

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:17:35 ; Search time 42 Seconds
(without alignments)
907.187 Million cell updates/sec

Title: US-10-783-415-4
Perfect score: 396
Sequence: 1 MVDFAMDVYKNLYSDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: OLIGOL
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	396	2 I75615	mammary tumor inte
2	15	3.8	418	2 T45807	translation initia
3	8	2.0	136	2 PQ0517	pol protein - ovin
4	8	2.0	393	2 H81745	conserved hypothet
5	8	2.0	425	2 S32866	outL protein - Erw
6	8	2.0	432	2 T33118	hypothetical prote
7	8	2.0	436	2 B86469	protein F12K21.16
8	8	2.0	4092	1 S38128	dyein heavy chain
9	7	1.8	65	2 B86846	hypothetical prote
10	7	1.8	98	2 B64097	hypothetical prote
11	7	1.8	123	1 T27492	cytochrome c 2C116
12	7	1.8	141	2 C24338	hemoglobin alpha-T
13	7	1.8	168	2 E81749	conserved hypothet
14	7	1.8	172	2 S56905	hypothetical prote
15	7	1.8	179	2 S05440	alpha-2u-globulin
16	7	1.8	181	1 UART	alpha-2u-globulin
17	7	1.8	181	2 I52504	alpha-2u-globulin
18	7	1.8	181	2 S05244	alpha-2u-globulin
19	7	1.8	181	2 I65319	alpha-2u-globulin
20	7	1.8	185	2 S75549	hypothetical prote
21	7	1.8	193	2 H83794	hypothetical prote
22	7	1.8	202	2 A83777	hypothetical prote
23	7	1.8	222	2 F70236	conserved hypothet
24	7	1.8	231	2 C69041	hypothetical prote
25	7	1.8	246	2 AG3644	flagellar biosynth
26	7	1.8	252	2 F90513	conserved hypothet
27	7	1.8	290	2 E83914	transcription regu
28	7	1.8	307	2 T01156	hypothetical prote
29	7	1.8	313	2 A34785	DNA-binding protei

30	7	1.8	332	2 JC5465	2,3-dihydroxybiphe
31	7	1.8	334	2 D85815	probable reductase
32	7	1.8	334	2 G64961	hypothetical prote
33	7	1.8	334	2 E90967	probable reductase
34	7	1.8	358	2 C42026	cyclic AMP respons
35	7	1.8	359	2 T34602	probable NTP prop
36	7	1.8	365	2 C69005	succinyl-CoA synth
37	7	1.8	367	2 T36116	probable oxidoredu
38	7	1.8	369	2 I78877	CAMP responsive el
39	7	1.8	375	2 JC8029	ornithine carbamoy
40	7	1.8	389	1 A39429	CAMP response elem
41	7	1.8	389	2 T47323	hypothetical prote
42	7	1.8	395	2 B71244	hypothetical prote
43	7	1.8	396	2 C75209	hypothetical prote
44	7	1.8	400	2 T33073	hypothetical prote
45	7	1.8	402	2 AD0148	tyrosine-specific

ALIGNMENTS

RESULT 1

I75615
mammary tumor integration site 6 oncogene protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999
C:Accession: I75615; I75614
R:Marchetti, A.; Buttitta, F.; Miyazaki, S.; Gallahan, D.; Smith, G.H.; Callahan, R.
J. Virol. 69, 1932-1938, 1995
A>Title: Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary
A:Reference number: I56914; MUID:95156630; PMID:7853537
A:Accession: I75615
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-396 <MAR1>
A:Cross-references: GB:S75224; NID:g913340; PID:g913341
A:Accession: I75614
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 235-268, 'LKIKFPQSKGIRYVNOAVSGFMEIHRAAVGRPEGHRLGRDPLENG' <MAR2>
A:Cross-references: GB:S75223; NID:g913338; PIDN:AAC00047.1; PID:g913339
A>Note: hypothetical mutant fusion protein
C:Genetics:
A:Gene: Int-6

Query Match 100.0%; Score 396; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAMDVYKNLYSDIPHALEKETTVAQLKQLQAEPIVKMFEDPETTRQMOSTRD 60
DB 1 MVDFAMDVYKNLYSDIPHALEKETTVAQLKQLQAEPIVKMFEDPETTRQMOSTRD 60

QY 61 GRMLFDYLADKHGFRQEVLDLTRYAKFOYECNYSAGAEYLYFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYLADKHGFRQEVLDLTRYAKFOYECNYSAGAEYLYFFRVLVPATDRNALSSL 120

QY 121 MGKLAISEILMQNWDAAEMDLTLKKTETIDNNSSVSSPLQSLQORTLTIHWSLFFVFNHPKGR 180
DB 121 MGKLAISEILMQNWDAAEMDLTLKKTETIDNNSSVSSPLQSLQORTLTIHWSLFFVFNHPKGR 180

QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITKDKVRKQVLDLKVIVQESYTYK 240
DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITKDKVRKQVLDLKVIVQESYTYK 240

QY 241 DPITEFVCLYVPFDGAKKRECEVLNDDFLVACLEDFIENARLFIETFCRIHQ 300
DB 241 DPITEFVCLYVPFDGAKKRECEVLNDDFLVACLEDFIENARLFIETFCRIHQ 300

QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
DB 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360

QY 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
 |||||
 Db 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 2

T45807
 translation initiation factor 3-like protein - Arabidopsis thaliana
 N:Alternate names: protein F2809.140
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
 C:Accession: T45807
 R:Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23014
 A:Accession: T45807
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-418 <BEN>
 A:Cross-references: UNIPROT:Q9M2L8; EMBL:AL137080
 A:Experimental source: cultivar Columbia; BAC clone F2809
 C:Genetics:
 A:Map position: 3
 A:Introns: 74/1; 125/3; 201/3; 219/2; 292/3; 340/2; 386/2
 A:Note: F2809.140

Query Match 3.8%; Score 15; DB 2; Length 418;
 Best Local Similarity 100.0%; Pred. No. 4.1e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 LMGKLASEILMQNWD 134
 |||||
 Db 171 LMGKLASEILMQNWD 185

RESULT 3

PQ0517
 pol protein - ovine lentivirus OLV-CV1 (strain CUI) (fragment)
 C:Species: ovine lentivirus OLV-CV1
 C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
 C:Accession: PQ0517
 R:Campbell, B.G.; Thompson, D.R.; Williams, J.R.; Campbell, S.G.; Avery, R.J.
 J. Gen. Virol. 74, 201-210, 1993
 A:Title: Characterization of a New York ovine lentivirus isolate.
 A:Reference number: PQ0517; MUID:93155645; PMID:8381461
 A:Accession: PQ0517
 A:Molecule type: DNA
 A:Residues: 1-136 <CAM>
 A:Cross-references: UNIPROT:Q7LZ00
 C:Genetics:
 A:Gene: pol
 C:Superfamily: pol polyprotein
 C:Keywords: polyprotein

Query Match 2.0%; Score 8; DB 2; Length 136;
 Best Local Similarity 100.0%; Pred. No. 2.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 VITNKDVR 219
 |||||
 Db 118 VITNKDVR 125

RESULT 4

HB1745
 conserved hypothetical protein TC0058 [imported] - Chlamydia muridarum (strain Nigg)
 C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C:Accession: HB1745
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935
 A:Accession: H81745
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-393 <TET>
 A:Cross-references: UNIPROT:Q9PLP1; GB:AE002273; GB:AE002160; NID:g7190092; PIDN:AAF3894
 A:Experimental source: strain Nigg (MoPn)
 C:Genetics:
 A:Gene: TC0058

Query Match 2.0%; Score 8; DB 2; Length 393;
 Best Local Similarity 100.0%; Pred. No. 7.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 ALSSLWGK 123
 |||||
 Db 19 ALSSLWGK 26

RESULT 5

S32866
 outL protein - Erwinia carotovora
 C:Species: Erwinia carotovora
 C>Date: 08-Dec-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
 C:Accession: S32866; S31755
 R:Reeves, P.J.; Whitcombe, D.; Wharham, S.; Gibson, M.; Allison, G.; Bunce, N.; Barallon,
 Mol. Microbiol. 8, 443-456, 1993
 A:Title: Molecular cloning and characterization of 13 out genes from Erwinia carotovora
 bacteria.
 A:Reference number: S32857; MUID:93316842; PMID:8326859
 A:Accession: S32866
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-425 <REE>
 A:Cross-references: UNIPROT:P31708; EMBL:X70049; NID:g42184; PIDN:CAA49653.1; PID:g42194
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, /./
 C:Genetics:
 A:Gene: outL
 C:Keywords: transmembrane protein

Query Match 2.0%; Score 8; DB 2; Length 425;
 Best Local Similarity 100.0%; Pred. No. 7.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RVLVPATD 113
 |||||
 Db 76 RVLVPATD 83

RESULT 6

T33118
 hypothetical protein B0511.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T33118
 R:Tin-Wollam, A.; Sutterer, C.; Ozersky, P.
 submitted to the EMBL Data Library, May 1998
 A:Description: The sequence of C. elegans cosmid B0511.
 A:Reference number: Z21285
 A:Accession: T33118
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-432 <TIN>
 A:Cross-references: UNIPROT:O61820; EMBL:AF067608; PIDN:AAAC17647.1; GSPDB:GN000019; CESP

A:Map position: 1
 A:Introns: 30/3; 122/3; 281/3

Query Match 2.0%; Score 8; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 7.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 GKLAISEIL 129
 Db 171 GKLAISEIL 178

RESULT 7
 B86469
 Protein F12K21.16 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: B86469
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.R.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B86469
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-436 <STO>
 A:Cross-references: UNIPROT:Q9LNL2; GB:AE005172; NID:98778261; PIDN:AAF79270.1; GSPDB:GN
 C:Genetics:
 A:Gene: F12K21.16
 A:Map position: 1
 C:Superfamily: Arabidopsis hypothetical protein F7N22.18

Query Match 2.0%; Score 8; DB 2; Length 436;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 LPDYLDK 71
 Db 372 LPDYLDK 379

RESULT 8
 S38128
 dynein heavy chain, cytosolic - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YKR054c
 N:Contains: dynein ATPase (EC 3.6.4.2)
 C:Species: Saccharomyces cerevisiae
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: S38128; S38130; S37701; S43936; S43077
 R:Visser, S.; Urrestazu, L.A.; Jauniaux, J.C.
 submitted to the Protein Sequence Database, March 1994
 A:Reference number: S38118
 A:Accession: S38128
 A:Molecule type: DNA
 A:Residues: 1-4092 <VIS>
 A:Cross-references: UNIPROT:P36022; EMBL:Z28279; NID:9486510; PIDN:CAA82132.1; PID:94865
 A:Experimental source: strain S288C
 R:van Vliet-Reedijk, J.C.; Planta, R.J.
 submitted to the Protein Sequence Database, March 1994
 A:Reference number: S38130
 A:Accession: S38130
 A:Molecule type: DNA
 A:Residues: 1-787 <VAN>
 A:Cross-references: EMBL:Z28279; GSPDB:GN00011; MIPS:YKR054c
 R:Li, Y.Y.; Yeh, E.Y.; Hays, T.; Bloom, K.S.
 submitted to the EMBL Data Library, May 1993
 A:Description: Disruption of mitotic spindle orientation in a yeast dynein mutant.
 A:Reference number: S37701
 A:Accession: S37701
 A:Molecule type: DNA
 A:Residues: 1-588 'C', 590-600, 'A', 602-1363, 'A', 1365-2631, 'P', 2633-2657, 2659, 'IGW', 2660-2
 A:Cross-references: EMBL:L15626

R:Eshel, D.; Urrestazu, L.A.; Visers, S.; Jauniaux, J.C.; van Vliet-Reedijk, J.C.; R.P.
 Proc. Natl. Acad. Sci. U.S.A. 90, 11172-11176, 1993
 A:Title: Cytoplasmic dynein is required for normal nuclear segregation in yeast.
 A:Reference number: S43936; MUID:94068566; PMID:8248224
 A:Accession: S43936
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-2117, 'IV', 2120-4092 <ESH>
 A:Cross-references: EMBL:Z21877; NID:9439287; PIDN:CAA79923.1; PID:9439288
 C:Genetics:
 A:Gene: SGD:DYN1; DHCl; MIPS:YKR054c
 A:Cross-references: SGD:S0001762; MIPS:YKR054c
 A:Map position: 11R
 C:Function:
 A:Description: probably acts in cytoplasmic microtubule-based motile processes includin
 red for nuclear segregation; required to achieve and maintain proper spindle positionin
 C:Superfamily: dynein heavy chain, cytosolic
 C:Keywords: ATP; heterotrimer; hydrolase; microtubule binding; nucleotide binding; P-1
 F:1796-1803/Region: nucleotide-binding motif A (P-loop)
 F:2074-2081/Region: nucleotide-binding motif A (P-loop)
 F:2418-2425/Region: nucleotide-binding motif A (P-loop)
 F:2760-2767/Region: nucleotide-binding motif A (P-loop)
 F:1802/Binding site: ATP (Lys) #status predicted
 F:2080/Binding site: ATP (Lys) #status predicted
 F:2424/Binding site: ATP (Lys) #status predicted
 F:2766/Binding site: ATP (Lys) #status predicted

Query Match 2.0%; Score 8; DB 1; Length 4092;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 AMNIEKKL 377
 Db 3594 AMNIEKKL 3601

RESULT 9
 B86846
 hypothetical protein yscD [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
 C:Accession: B86846
 R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli,
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
 A:Reference number: A86625; MUID:21235186; PMID:11337471
 A:Accession: B86846
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-65 <STO>
 A:Cross-references: UNIPROT:Q9CER5; GB:AE005176; PID:gl2724793; PIDN:AAK05868.1; GSPDB:
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: yscD

Query Match 1.8%; Score 7; DB 2; Length 65;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 IIDLFLY 189
 Db 51 IIDLFLY 57

RESULT 10
 B64097
 hypothetical protein HI0828 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
 C:Accession: B64097
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, P.
 ; Goayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: B64097
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-98 <TIGR>
A:Cross-references: UNIPROT:P44887; GB:U32765; GB:L42023; NID:g1573838; PIDN:AAC22486.1;

Query Match 1.8%; Score 7; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LKQLQAE 38
|||||
DB 28 LKQLQAE 34

RESULT 11
T27492
Cytochrome c ZC116.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 03-Nov-2000 #sequence_revision 03-Nov-2000 #text_change 09-Jul-2004
C:Accession: T27492
R:Smyle, R.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z20376
A:Accession: T27492
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-123 <WIL>
A:Cross-references: UNIPROT:Q23240; EMBL:Z74046; PIDN:CAA98555.1; GSPDB:GN00023; CESP:ZC
A:Experimental source: clone ZC116
C:Genetics:
A:Gene: CESP:ZC116.2
A:Map position: 5
A:Introns: 24/3; 94/3
C:Superfamily: cytochrome c; cytochrome c homology
C:Keywords: chromoprotein; heme; iron; metalloprotein
F:20-113/Domain: cytochrome c homology <CYC>
F:30,33/Binding site: heme (Cys) (covalent) #status predicted
F:34,95/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Query Match 1.8%; Score 7; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 LFDYLAD 70
|||||
DB 79 LFDYLAD 85

RESULT 12
C24338
hemoglobin alpha-T5 chain - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 08-Aug-1987 #sequence_revision 08-Aug-1987 #text_change 12-Jul-2004
C:Accession: C24338
R:Banville, D.; Williams, J.G.
Nucleic Acids Res. 13, 5407-5421, 1985
A:Title: The pattern of expression of the Xenopus laevis tadpole alpha-globin genes and
A:Reference number: A93578; MUID:85297748; PMID:2993998
A:Accession: C24338
A:Molecule type: mRNA
A:Residues: 1-141 <BAN>
A:Cross-references: UNIPROT:P06638; GB:X02798; NID:G64542; PIDN:CAA26566.1; PID:G64543
C:Comment: The initiator Met is not shown.
C:Superfamily: globin; globin homology
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxygen carrier
F:2-141/Domain: Globin homology <GUB>
F:87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 1.8%; Score 7; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 DDIPHAL 21
|||||
DB 74 DDIPHAL 80

RESULT 13
E81749
conserved hypothetical protein TC0041 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: E81749
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: E81749
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-168 <TET>
A:Cross-references: UNIPROT:Q9PLQ4; GB:AE002271; GB:AE002160; NID:g7190063; PIDN:AAF3893
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0041

Query Match 1.8%; Score 7; DB 2; Length 168;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 KLRCEES 268
|||||
DB 37 KLRCEES 43

RESULT 14
S56905
hypothetical protein YJL124c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein J0714
C:Species: Saccharomyces cerevisiae
C:Date: 05-May-1995 #sequence_revision 08-Sep-1995 #text_change 09-Jul-2004
C:Accession: S56905
R:Cziepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56891
A:Accession: S56905
A:Molecule type: DNA
A:Residues: 1-172 <CZI>
A:Cross-references: UNIPROT:P47017; EMBL:Z49399; NID:g1008319; PIDN:CAA89419.1; PID:g100
C:Genetics:
A:Gene: SGD:LSM1; MIPS:YJL124c
A:Cross-references: SGD:S0003660
A:Map position: 10L
C:Superfamily: hypothetical protein YJL124c

Query Match 1.8%; Score 7; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 RDGRMLF 65
|||||
DB 59 RDGRMLF 65

RESULT 15
S05440
alpha-2u-globulin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004
C:Accession: S05440; S70347

R;Gao, F.; Endo, H.; Yamamoto, M.
 Nucleic Acids Res. 17, 4629-4636, 1989
 A;Title: Length heterogeneity in rat salivary gland alpha-2-mu globulin mRNAs: multiple
 A;Reference number: S05440; MUID:89315206; PMID:2473438
 A;Accession: S05440
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-179 <GAO>
 A;Cross-references: UNIPROT:063213; EMBL:X14552; NID:G55569; PIDN:CAA32690.1; PID:G55570
 A;Experimental source: salivary gland
 R;Bayard, C.; Holmquist, L.; Vesterberg, O.
 Biochim. Biophys. Acta 1290, 129-134, 1996
 A;Title: Purification and identification of allergenic alpha(2u)-globulin species of rat
 A;Reference number: S70347; MUID:96254071; PMID:8645715
 A;Accession: S70347
 A;Molecule type: protein
 A;Residues: 20-44 <BAY>
 A;Experimental source: urine
 C;Function:
 A;Description: transport protein of hydrophobic molecules
 A;Note: eight-stranded beta-barrel conformation; interior cavity lined with hydrophobic
 C;Superfamily: lipocalin; lipocalin homology
 C;Keywords: glycoprotein; lipid binding; salivary gland; urine
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;20-179/Product: alpha-2u-globulin #status experimental <MAT>
 F;29-176/Domain: lipocalin homology <LIP>
 F;54/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.8%; Score 7; DB 2; Length 179;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 RDNIIDL 186
 |||||
 Db 164 RDNIIDL 170

Search completed: March 8, 2005, 20:27:22
 Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.


OM protein - protein search, using sw model

Run on: March 8, 2005, 20:08:14 ; Search time 165 Seconds
(without alignments)
928.225 Million cell updates/sec

Title: US-10-783-415-4

Perfect score: 396

Sequence: 1 MVDFAMDVYKNLYSDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table:  Gapex 60.0, Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: A Geneseq 16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	100.0	396	2 AAW02113	Aaw02113 Murine In
2	396	100.0	396	2 AAW02112	Aaw02112 Human hom
3	396	100.0	396	2 AAB47920	Aab47920 Murine In
4	396	100.0	396	2 AAB47921	Aab47921 Human Int
5	396	100.0	396	6 ABU05167	Abu05167 Human exp
6	396	100.0	396	6 ABU05169	Abu05169 Human exp
7	396	100.0	396	6 ABU05158	Abu05158 Human exp
8	396	100.0	396	7 ADJ70258	Adj70258 Human hea
9	396	100.0	396	8 ADN30996	Adn30996 Human Int
10	396	100.0	396	8 ADN30994	Adn30994 Murine In
11	396	100.0	445	5 ABB57306	Abb57306 Mouse isc
12	396	100.0	445	6 ABR39935	Abr39935 Human pro
13	396	100.0	445	6 ABU05160	Abu05160 Human exp
14	396	100.0	445	6 ABU05163	Abu05163 Human exp
15	396	100.0	445	6 ABU05157	Abu05157 Human exp
16	396	100.0	445	6 ABU05166	Abu05166 Human exp
17	396	100.0	445	6 ABU05159	Abu05159 Human exp
18	396	100.0	445	6 ABU05162	Abu05162 Human exp
19	396	100.0	445	6 ABU05164	Abu05164 Human exp
20	396	100.0	445	6 ABU05168	Abu05168 Human exp
21	396	100.0	445	7 ABM85381	Abm85381 Human pro
22	396	100.0	445	8 ADL13112	Adl13112 Human ste
23	396	100.0	445	8 ABM81005	Abm81005 Tumour-as
24	295	74.5	445	6 ABU05161	Abu05161 Human exp
25	204	51.5	320	7 ABM85380	Abm85380 Mouse pro

26	166	41.9	445	6	ABU05165	Abu05165 Human exp
27	103	26.0	117	5	ABP43106	Abp43106 Human ova
28	91	23.0	155	4	ABG21242	Abg21242 Novel hum
29	45	11.4	45	4	AAM22004	Aam22004 Peptide #
30	45	11.4	45	4	ABB44381	Abb44381 Peptide #
31	45	11.4	45	4	AAM38359	Aam38359 Peptide #
32	45	11.4	45	4	ABB27240	Abb27240 Protein #
33	45	11.4	45	4	AAM78125	Aam78125 Human bon
34	45	11.4	45	4	AAM65454	Aam65454 Human bra
35	45	11.4	45	4	ABG59761	Abg59761 Human liv
36	45	11.4	45	5	ABG47142	Abg47142 Human pep
37	43	10.9	1202	4	ABG21244	Abg21244 Novel hum
38	43	10.9	1202	7	ADF60389	Adf60389 Human con
39	42	10.6	42	4	AAM21570	Aam21570 Peptide #
40	42	10.6	42	4	ABB41600	Abb41600 Peptide #
41	42	10.6	42	4	ABB43923	Abb43923 Peptide #
42	42	10.6	42	4	AAM37849	Aam37849 Peptide #
43	42	10.6	42	4	AAM35393	Aam35393 Peptide #
44	42	10.6	42	4	ABB25431	Abb25431 Protein #
45	42	10.6	42	4	ABB26840	Abb26840 Protein #

ALIGNMENTS

RESULT 1

AAW02113

ID AAW02113 standard; protein; 396 AA.

XX AAW02113;

DT 15-MAY-1997 (first entry)

DE Murine Int6 protein associated with MMTV integration and tumour growth.

KW MMTV; mouse mammary tumour virus; Int6; breast cancer; neoplasia;

KW diagnosis; treatment; immunotherapy; vaccine; probe; primer.

XX Mus musculus.

Key Location/Qualifiers

Misc-difference 27 /note= "potential cAMP/CGMP-dependent protein kinase phosphorylation site"

Misc-difference 51 /note= "potential protein kinase C phosphorylation site"

Misc-difference 57 /note= "potential casein kinase II phosphorylation site"

Misc-difference 90 /note= "potential tyrosine kinase phosphorylation site"

Misc-difference 94 /note= "potential glycosylation site"

Misc-difference 112 /note= "potential protein kinase C phosphorylation site"

Misc-difference 149 /note= "potential glycosylation site"

Misc-difference 214 /note= "potential casein kinase II phosphorylation site"

Misc-difference 238 /note= "potential casein kinase II phosphorylation site"

Misc-difference 314 /note= "potential casein kinase II phosphorylation site"

Misc-difference 363 /note= "potential protein kinase C phosphorylation site"

Misc-difference 381 /note= "potential casein kinase II phosphorylation site"

XX WO9624672-A1.

XX 15-AUG-1996.

XX 09-FEB-1996; 96WO-US001884.

PR 09-FEB-1995; 95US-00385998.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Marchetti A, Buttitta F, Smith GH, Callahan R;
 XX WPI; 1996-384444/38.
 DR N-PSDB; AAT36177.
 XX DNA encoding Int6 tumour associated protein - and use of reagents derived
 PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
 PS Disclosure; Page 60-61; 93pp; English.
 XX
 CC AAW02113 is the murine Int6 protein. The Int6 gene is located at
 CC chromosome 15 of a mouse genome. The Int6 gene is associated with MMTV
 CC (mouse mammary tumour virus) integration into a host genome during
 CC tumorigenesis. Primers and probes may be derived from the Int6 gene
 CC sequence and used for detection in assays to diagnose MMTV infection, or
 CC any other Int6 gene integration. Antibodies against the Int6 protein can
 CC be used in the same way. The DNA and protein may also be used gene
 CC therapy and vaccines
 XX
 SQ Sequence 396 AA;
 Query Match 100.0%; Score 396; DB 2; Length 396;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVDFAVDVYKNLYSDDIPHALREKRTTVAQLKQLAETETPIVKMFEDPETTRQMSTRD 60
 DB 1 MVDFAVDVYKNLYSDDIPHALREKRTTVAQLKQLAETETPIVKMFEDPETTRQMSTRD 60
 QY 61 GRMLFDYLDKKGFRQYDLTLRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
 DB 61 GRMLFDYLDKKGFRQYDLTLRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
 QY 121 WGLASEILMOWNDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
 DB 121 WGLASEILMOWNDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
 QY 181 DNIIDFLYQOYLNAIQTCMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
 DB 181 DNIIDFLYQOYLNAIQTCMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
 QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
 DB 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
 QY 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIEKTK 360
 DB 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
 RESULT 2
 AAW02112
 ID AAW02112 standard; protein; 396 AA.
 XX
 AC AAW02112;
 XX
 DT 14-MAY-1997 (first entry)
 XX Human homologue of Int6 protein associated with MMTV integration.
 DE
 XX MMTV; mouse mammary tumour virus; Int6; breast cancer; neoplasia;
 KW diagnosis; treatment; immunotherapy; vaccine; probe; primer.
 XX Homo sapiens.
 OS
 XX

PN WO9624672-A1.
 XX 15-AUG-1996.
 XX 09-FEB-1996; 96WO-US001884.
 XX 09-FEB-1995; 95US-00385998.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Marchetti A, Buttitta F, Smith GH, Callahan R;
 PI WPI; 1996-384444/38.
 XX N-PSDB; AAT36148.
 DR
 DR DNA encoding Int6 tumour associated protein - and use of reagents derived
 PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
 XX Claim 3; Page 62-63; 93pp; English.
 PS
 CC AAW02112 represents the human homologue of murine Int6 protein. The Int6
 CC gene is located at chromosome 15 of the mouse genome and is associated
 CC with MMTV (mouse mammary tumour virus) integration into a host genome
 CC during tumorigenesis. Primers and probes may be derived from the Int6
 CC cDNA sequence and used for the detection of the Int6 gene. These can be
 CC used in assays to diagnose MMTV infection, or any other Int6 gene
 CC integration. Antibodies against the Int6 protein can be used in the same
 CC way
 XX
 SQ Sequence 396 AA;
 Query Match 100.0%; Score 396; DB 2; Length 396;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVDFAVDVYKNLYSDDIPHALREKRTTVAQLKQLAETETPIVKMFEDPETTRQMSTRD 60
 DB 1 MVDFAVDVYKNLYSDDIPHALREKRTTVAQLKQLAETETPIVKMFEDPETTRQMSTRD 60
 QY 61 GRMLFDYLDKKGFRQYDLTLRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
 DB 61 GRMLFDYLDKKGFRQYDLTLRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
 QY 121 WGLASEILMOWNDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
 DB 121 WGLASEILMOWNDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
 QY 181 DNIIDFLYQOYLNAIQTCMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
 DB 181 DNIIDFLYQOYLNAIQTCMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
 QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
 DB 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
 QY 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIEKTK 360
 DB 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
 RESULT 3
 AAW02112
 ID AAW02112 standard; protein; 396 AA.
 XX
 AC AAW02112;
 XX
 DT 21-MAY-2002 (first entry)
 XX Murine Int6.
 DE

XX Murine; human; Int6; integration site; deregulation; neoplasia;
 KW mouse mammary tumour virus; MMTV; cancer; immunotherapy; gene therapy;
 KW prenatal screening; foetus; vaccine.
 XX Mus musculus.

XX Location/Qualifiers
 XX Key
 XX Modified-site 27 /note= "Phosphorylation site for cAMP/cGMP-dependent
 FT protein kinase"
 FT Modified-site 51
 FT Modified-site 57 /note= "Phosphorylation site for protein kinase C"
 FT Modified-site 57 /note= "Phosphorylation site for casein kinase II"
 FT Modified-site 90 /note= "Phosphorylation site for tyrosine kinase"
 FT Modified-site 94 /note= "Phosphorylation site for tyrosine kinase and
 FT glycosylation site"
 FT Modified-site 112 /note= "Phosphorylation site for protein kinase C"
 FT Modified-site 149 /note= "Phosphorylation site for tyrosine kinase and
 FT glycosylation site"
 FT Modified-site 214 /note= "Phosphorylation site for casein kinase II"
 FT Modified-site 238 /note= "Phosphorylation site for casein kinase II"
 FT Modified-site 314 /note= "Phosphorylation site for casein kinase II"
 FT Modified-site 363 /note= "Phosphorylation site for protein kinase C"
 FT Modified-site 381 /note= "Phosphorylation site for casein kinase II"
 FT
 XX US6342392-B1.
 XX 29-JAN-2002.
 XX 23-AUG-1999; 99US-00378842.
 XX 09-FEB-1995; 95US-00385998.
 XX 09-FEB-1996; 96WO-US001884.
 XX 25-SEP-1997; 97US-00875847.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Marchetti A, Buttitta F, Smith GH, Callahan R;
 XX WPI; 1996-384444/38.
 XX N-PSDB; AAI72499.
 XX DNA encoding Int6 tumour associated protein - and use of reagents derived
 PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
 XX Disclosure; Col 35-38; 45pp; English.

XX This sequence shows murine Int6. Int6 is an integration site for mouse
 CC mammary tumour virus (MMTV), which causes deregulation of expression of
 CC cellular genes adjacent to the site of MMTV integration in mammary
 CC tumours. The Int6 protein has been found to be highly conserved across
 CC species, with Drosophila Int6 being 60% identical to human/mouse Int6.
 CC This indicates that Int6 is serving a basic life function. The method of
 CC the invention comprises assaying a sample to detect a human Int6 nucleic
 CC acid sequence, or its fragment, by contacting the sample with a sequence
 CC of at least 15 consecutive nucleotides of human Int6 cDNA or a
 CC conservative variant of it, where a disrupted expression or loss of
 CC expression of the variant is associated with neoplasia. The method is
 CC useful for prenatal screening of a foetus or to pre-symptomatically
 CC screen a subject at risk of having cancer. Detecting mutations in the
 CC Int6 gene can provide diagnostic and prognostic information. The nucleic
 CC acids and proteins are useful in immunotherapy, gene therapy or as

CC vaccines for treating or preventing cancer. The nucleic acids are useful
 CC as probes for isolating homologues of Int6 gene or for detecting
 CC mutations in the Int6 gene
 XX
 SQ Sequence 396 AA;
 Query Match 100.0%; Score 396; DB 2; Length 396;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVDFAVDVYKNLYSDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMQSTRD 60
 DB 1 MVDFAVDVYKNLYSDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMQSTRD 60
 QY 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLFFRVLVPAATDRNALSSL 120
 DB 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLFFRVLVPAATDRNALSSL 120
 QY 121 WGLASEILMKNWDAAMEDLTRKETIDNNSVSPLOQRTWLIIHWSLFFVFNHPKGR 180
 DB 121 WGLASEILMKNWDAAMEDLTRKETIDNNSVSPLOQRTWLIIHWSLFFVFNHPKGR 180
 QY 181 DNIIDFLYQPYQLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTK 240
 DB 181 DNIIDFLYQPYQLNAIQTMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTK 240
 QY 241 DPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFTIENARLFIPTFCRIHQ 300
 DB 241 DPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFTIENARLFIPTFCRIHQ 300
 QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVGNNNAVSPYQVIEKTK 360
 DB 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVGNNNAVSPYQVIEKTK 360
 QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
 DB 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 4
 AAB47921
 ID AAB47921 standard; protein; 396 AA.
 AC AAB47921;
 XX 21-MAY-2002 (first entry)
 XX Human Int6.
 XX Murine; human; Int6; integration site; deregulation; neoplasia;
 KW mouse mammary tumour virus; MMTV; cancer; immunotherapy; gene therapy;
 KW prenatal screening; foetus; vaccine; chromosome 8q22-q24.
 XX Homo sapiens.
 XX US6342392-B1.
 XX 29-JAN-2002.
 XX 23-AUG-1999; 99US-00378842.
 XX 09-FEB-1995; 95US-00385998.
 XX 09-FEB-1996; 96WO-US001884.
 XX 25-SEP-1997; 97US-00875847.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Marchetti A, Buttitta F, Smith GH, Callahan R;
 XX WPI; 1996-384444/38.
 XX N-PSDB; AAI72499.
 XX DNA encoding Int6 tumour associated protein - and use of reagents derived

PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
PS Disclosure; Col 39-44; 45pp; English.
XX
XX This sequence shows human Int6. The human Int6 coding sequence was
CC isolated by using murine Int6 sequences as probes. Human Int6 is
CC organised into 13 exons as is murine Int6, and contains a CA-repeat in
CC the 7th intron. Human Int6 has been localised to chromosome 8, more
CC specifically to 8q22-q24. Int6 is an integration site for mouse mammary
CC tumour virus (MMTV), which causes deregulation of expression of cellular
CC genes adjacent to the site of MMTV integration in mammary tumours. The
CC Int6 protein has been found to be highly conserved across species, with
CC Drosophila Int6 being 60% identical to human/mouse Int6. This indicates
CC that Int6 is serving a basic life function. The method of the invention
CC comprises assaying a sample to detect a human Int6 nucleic acid sequence,
CC or its fragment, by contacting the sample with a sequence of at least 15
CC consecutive nucleotides of human Int6 cDNA or a conservative variant of
CC it, where a disrupted expression or loss of expression of the variant is
CC associated with neoplasia. The method is useful for prenatal screening of
CC a foetus or to pre-symptomatically screen a subject at risk of having
CC cancer. Detecting mutations in the Int6 gene can provide diagnostic and
CC prognostic information. The nucleic acids and proteins are useful in
CC immunotherapy, gene therapy or as vaccines for treating or preventing
CC cancer. The nucleic acids are useful as probes for isolating homologues
CC of Int6 gene or for detecting mutations in the Int6 gene
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 396; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVKNLYSDDI PHALREKRTTVVAQLKQLAQETPIVKMPEDPETTRQMSTRD 60
DB 1 MVDFAVDVKNLYSDDI PHALREKRTTVVAQLKQLAQETPIVKMPEDPETTRQMSTRD 60

QY 61 GRMLFDYLDKKGFRQYLDLTRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYLDKKGFRQYLDLTRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120

QY 121 WGLASEILMKNWDAAMEDLTRKETIDNNSVSSPLQSQOORTWLHWSLFFVFNHPKGR 180
DB 121 WGLASEILMKNWDAAMEDLTRKETIDNNSVSSPLQSQOORTWLHWSLFFVFNHPKGR 180

QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDKRROVLKDLVKVIOQESYTYK 240
DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDKRROVLKDLVKVIOQESYTYK 240

QY 241 DPITEFVECLYVNFDFDGAQKKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300

QY 301 CISINMLADKNMTPAEARWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETKX 360
DB 301 CISINMLADKNMTPAEARWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETKX 360

QY 361 SLFSRQMLAMNIEKKLNQSRSEAPNWTQDSGFY 396
DB 361 SLFSRQMLAMNIEKKLNQSRSEAPNWTQDSGFY 396

RESULT 5
ABU05167
ID ABU05167 standard; protein; 396 AA.
XX
AC ABU05167;
XX
DT 29-JAN-2003 (first entry)
DE Human expressed protein tag (EPT) #1833.
XX
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
DR
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1833; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 396; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVKNLYSDDI PHALREKRTTVVAQLKQLAQETPIVKMPEDPETTRQMSTRD 60
DB 1 MVDFAVDVKNLYSDDI PHALREKRTTVVAQLKQLAQETPIVKMPEDPETTRQMSTRD 60

QY 61 GRMLFDYLDKKGFRQYLDLTRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYLDKKGFRQYLDLTRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120

QY 121 WGLASEILMKNWDAAMEDLTRKETIDNNSVSSPLQSQOORTWLHWSLFFVFNHPKGR 180
DB 121 WGLASEILMKNWDAAMEDLTRKETIDNNSVSSPLQSQOORTWLHWSLFFVFNHPKGR 180

QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDKRROVLKDLVKVIOQESYTYK 240
DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDKRROVLKDLVKVIOQESYTYK 240

QY 241 DPITEFVECLYVNFDFDGAQKKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKKLRCESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300

```
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVGNNNAVSPYQVIEKTK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVGNNNAVSPYQVIEKTK 360
Qy 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

RESULT 6
ABU05169
ID ABU05169 standard; protein; 396 AA.
XX
AC ABU05169;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1835.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN W0200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1835; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
```

```
XX SQ Sequence 396 AA;
Query Match 100.0%; Score 396; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVDPAFMDVYKLYSDDIPHALREKRTTVVAQLKQLOQAETEPVVKMFEDPETTRQMQSTRD 60
Db 1 MVDPAFMDVYKLYSDDIPHALREKRTTVVAQLKQLOQAETEPVVKMFEDPETTRQMQSTRD 60
Qy 61 GRMLFDYLADKKGPRQYELDTLYRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
Db 61 GRMLFDYLADKKGPRQYELDTLYRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
Qy 121 WGKLASEILMNQWDAAMEDLTRKKTIDNNSVSPLOSQORTWLIIHWSLFPVFNHPKGR 180
Db 121 WGKLASEILMNQWDAAMEDLTRKKTIDNNSVSPLOSQORTWLIIHWSLFPVFNHPKGR 180
Qy 181 DNIIDFLYQOYVNAIQTMCPHILRYLTAVITNKDVRKROVLDLVKVIQOESYTYK 240
Db 181 DNIIDFLYQOYVNAIQTMCPHILRYLTAVITNKDVRKROVLDLVKVIQOESYTYK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVGNNNAVSPYQVIEKTK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVGNNNAVSPYQVIEKTK 360
Qy 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

RESULT 7
ABU05158
ID ABU05158 standard; protein; 396 AA.
XX
AC ABU05158;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1824.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN W0200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
```



```
QY 301 CINSINMLADKLNMTPEAERWIVNLIARNLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CINSINMLADKLNMTPEAERWIVNLIARNLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 9
ADN30996
ID ADN30996 standard; protein; 396 AA.
XX
AC ADN30996;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human Int6 protein.
XX
KW Human; Int6; mammary epithelial cellular growth; cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN US6737251-B2.
XX
PD 18-MAY-2004.
XX
PF 14-MAY-2001; 2001US-00858152.
XX
PR 09-FEB-1995; 95US-00385998.
XX
PR 09-FEB-1996; 96US-00875847.
XX
PR 09-FEB-1996; 96WO-US001884.
XX
PR 23-AUG-1999; 99US-00378842.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
DR WPI; 2004-387097/36.
XX
DR N-PSDB; ADN30995.
XX
PT Novel tumor Int6 recombinant protein that deregulates mammary epithelial
PT cellular growth, useful for treating cancer.
XX
PS Claim 1; SEQ ID NO 4; 44pp; English.
XX
CC The invention relates to the Int6 protein and the cDNA encoding it. The
CC Int6 protein deregulates mammary epithelial cellular growth. The cDNA and
CC protein are useful as vaccines for treating cancer. This sequence
CC represents the human Int6 protein of the invention.
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 396; DB 8; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLQAEETPIVGMFEDPETTRQMSTRD 60
Db 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLQAEETPIVGMFEDPETTRQMSTRD 60
QY 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPATDNLSSL 120
Db 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPATDNLSSL 120
QY 121 WGLKASILMNQNDAAEMEDLTRKETIDNNNSVSPLOSQLOORTWLIHWSLFVFNHPKGR 180
Db 121 WGLKASILMNQNDAAEMEDLTRKETIDNNNSVSPLOSQLOORTWLIHWSLFVFNHPKGR 180
QY 181 DNIIDFLYQPOYLNAIQTCWPHILRLYLTAVITNKDVRKRRQVLDLVKVIQOESYTK 240
Db 181 DNIIDFLYQPOYLNAIQTCWPHILRLYLTAVITNKDVRKRRQVLDLVKVIQOESYTK 240
```

```
QY 241 DPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
QY 301 CINSINMLADKLNMTPEAERWIVNLIARNLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CINSINMLADKLNMTPEAERWIVNLIARNLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 10
ADN30994
ID ADN30994 standard; protein; 396 AA.
XX
AC ADN30994;
XX
DT 29-JUL-2004 (first entry)
XX
DE Murine Int6 protein.
XX
KW Mouse; Int6; mammary epithelial cellular growth; cancer; cytostatic.
XX
OS Mus musculus.
XX
PN US6737251-B2.
XX
PD 18-MAY-2004.
XX
PF 14-MAY-2001; 2001US-00858152.
XX
PR 09-FEB-1995; 95US-00385998.
XX
PR 09-FEB-1996; 96US-00875847.
XX
PR 09-FEB-1996; 96WO-US001884.
XX
PR 23-AUG-1999; 99US-00378842.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
DR WPI; 2004-387097/36.
XX
DR N-PSDB; ADN30993.
XX
PT Novel tumor Int6 recombinant protein that deregulates mammary epithelial
PT cellular growth, useful for treating cancer.
XX
PS Example 4; SEQ ID NO 2; 44pp; English.
XX
CC The invention relates to the Int6 protein and the cDNA encoding it. The
CC Int6 protein deregulates mammary epithelial cellular growth. The cDNA and
CC protein are useful as vaccines for treating cancer. This sequence
CC represents the murine Int6 protein of the invention.
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 396; DB 8; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLQAEETPIVGMFEDPETTRQMSTRD 60
Db 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLQAEETPIVGMFEDPETTRQMSTRD 60
QY 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPATDNLSSL 120
Db 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPATDNLSSL 120
QY 121 WGLKASILMNQNDAAEMEDLTRKETIDNNNSVSPLOSQLOORTWLIHWSLFVFNHPKGR 180
Db 121 WGLKASILMNQNDAAEMEDLTRKETIDNNNSVSPLOSQLOORTWLIHWSLFVFNHPKGR 180
```

QY 181 DNIIDFLYQOYLNAIQMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 240
 DB 181 DNIIDFLYQOYLNAIQMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 240
 QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
 DB 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
 QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOQVIEKTK 360
 DB 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOQVIEKTK 360
 QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
 DB 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 11
 ABB57306
 ID ABB57306 standard; protein; 445 AA.
 AC ABB57306;
 XX
 DT 07-MAR-2002 (first entry)
 XX
 DE Mouse ischaemic condition related protein sequence SEQ ID NO:856.
 XX
 XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
 XX
 OS Mus musculus.
 XX
 PN WO200188188-A2.
 XX
 PD 22-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-JP004192.
 XX
 PR 18-MAY-2000; 2000JP-00145977.
 XX
 XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 PA
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 XX
 DR WPI; 2002-034733/04.
 DR N-PSDB; AB199770.
 XX
 XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or by
 PT determining the expression profile of a gene group comprising these
 PT genes.
 XX
 XX Claim 2; Page 2125-2127; 2690pp; English.
 PS
 PS The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (AB199202 to AB199912, encoding the
 CC protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The expression
 CC levels or expression profiles produced by these genes are used as an
 CC indicator when screening for ischaemic condition-improving drugs or
 CC therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR
 CC primers for a mouse ischaemic condition related sequence, which are used
 CC in the exemplification of the present invention
 XX
 SQ Sequence 445 AA;

Query Match 100.0%; Score 396; DB 5; Length 445;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDFAMDVYKLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMOSTRD 60
 DB 50 MYDFAMDVYKLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFEDPETTRQMOSTRD 109
 QY 61 GRMLFDYLDADKHGFRQYLDTRYAKFOYECGNYSAAEYLYFFRVLVPATDARNALSSL 120
 DB 110 GRMLFDYLDADKHGFRQYLDTRYAKFOYECGNYSAAEYLYFFRVLVPATDARNALSSL 169
 QY 121 WGLASEIILMOWDAAMEDLTRKETIDNNSVSSPLOSLOORTWLIHWSLFVFFNHPKGR 180
 DB 170 WGLASEIILMOWDAAMEDLTRKETIDNNSVSSPLOSLOORTWLIHWSLFVFFNHPKGR 229
 QY 181 DNIIDFLYQOYLNAIQMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 240
 DB 230 DNIIDFLYQOYLNAIQMCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 289
 QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
 DB 290 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 349
 QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOQVIEKTK 360
 DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOQVIEKTK 409
 QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
 DB 410 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 12
 ABR39935
 ID ABR39935 standard; protein; 445 AA.
 XX
 AC ABR39935;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Human prostate selective polypeptide Pr327.
 XX
 KW Prostate; molecular marker; cancer; cytostatic; gene therapy; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003014298-A2.
 XX
 PD 20-FEB-2003.
 XX
 PF 02-AUG-2002; 2002WO-US024431.
 XX
 PR 03-AUG-2001; 2001US-0309470P.
 PR 30-OCT-2001; 2001US-0330747P.
 XX
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 PA
 XX Sun Z, Li X, Jay G, Kovacs KF, Shu Y, Pan W;
 PI WPI; 2003-256562/25.
 DR N-PSDB; ACC47340.
 XX
 XX New polynucleotide, useful for preparing a composition for treating
 PT prostate disease, e.g., cancer.
 PT
 XX Claim 5; Page 150-152; 212pp; English.
 XX
 CC The invention relates to prostate selective polynucleotides and
 CC polypeptides. The polynucleotides are expressed in prostate and are
 CC useful as molecular markers, as drug targets, and for detecting,
 CC monitoring, preventing or treating diseases and conditions related to
 CC prostate, such as prostate cancers. The present sequence represents a
 CC prostate specific polypeptide

```
XX SQ Sequence 445 AA;
Query Match 100.0%; Score 396; DB 6; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLQLOAETEPVVKMFEDPETTRQMOSTRD 60
DB 50 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLQLOAETEPVVKMFEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFOYECGNYSGAAEYLYFFRVLPATDNLSSL 120
DB 110 GRMLFDYLDKKGFRQBYLDTLYRYAKFOYECGNYSGAAEYLYFFRVLPATDNLSSL 169
QY 121 WGLASEILMNQNDAAEMEDTLRLKETIDNNSVSPLOSLOORTWLHWSLFFVFNHPKGR 180
DB 170 WGLASEILMNQNDAAEMEDTLRLKETIDNNSVSPLOSLOORTWLHWSLFFVFNHPKGR 229
QY 181 DNIIDLFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRRQVLDKLVKVIQESYTYK 240
DB 230 DNIIDLFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRRQVLDKLVKVIQESYTYK 289
QY 241 DPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
DB 290 DPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 349
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 13
ABU05160
ID ABU05160 standard; protein; 445 AA.
AC ABU05160;
XX
DT 29-JAN-2003 (first entry)
DE Human expressed protein tag (EPT) #1826.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
```

New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

Example 2; SEQ ID NO 1826; 134pp; English.

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, or protease inhibitor, a transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting a mammal an immunogenic response directed against any of the purified polypeptide, is useful for treating cancer, or the antibody that binds to this polypeptide. The polypeptide, or the antibody that binds to this polypeptide, is useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: this sequence does not appear in the printed or specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 445 AA;

Query Match 100.0%; Score 396; DB 6; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLQLOAETEPVVKMFEDPETTRQMOSTRD 60
DB 50 MVDFAVDVYKNLYSDDDIPHALREKRTTVVAQLQLOAETEPVVKMFEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFOYECGNYSGAAEYLYFFRVLPATDNLSSL 120
DB 110 GRMLFDYLDKKGFRQBYLDTLYRYAKFOYECGNYSGAAEYLYFFRVLPATDNLSSL 169
QY 121 WGLASEILMNQNDAAEMEDTLRLKETIDNNSVSPLOSLOORTWLHWSLFFVFNHPKGR 180
DB 170 WGLASEILMNQNDAAEMEDTLRLKETIDNNSVSPLOSLOORTWLHWSLFFVFNHPKGR 229
QY 181 DNIIDLFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRRQVLDKLVKVIQESYTYK 240
DB 230 DNIIDLFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRRQVLDKLVKVIQESYTYK 289
QY 241 DPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
DB 290 DPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 349
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSOMLAWNTEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 14
ABU05163
ID ABU05163 standard; protein; 445 AA.
XX
AC ABU05163;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1829.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 OS Homo sapiens.
 XX WO200278524-A2.
 PN 10-OCT-2002.
 XX 28-MAR-2002; 2002WO-US009671.
 XX 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX (ZYCO-) ZYCOS INC.
 XX Chicx RM, Tomlinson AJ, Urban RG;
 XX WPI; 2003-040607/03.
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX Example 2; SEQ ID NO 1829; 134pp; English.
 XX The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

Query Match
 Best Local Similarity 100.0%; Score 396; DB 6; Length 445;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVKNLYSDDIPHALREKRTTVVAQLKQLAQETPIVKMFEDPETTRQMOSTRD 60
 DB 50 MVDFAVDVKNLYSDDIPHALREKRTTVVAQLKQLAQETPIVKMFEDPETTRQMOSTRD 109
 QY 61 GRMLFDYLDKGFGRQYLDLTRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
 DB 110 GRMLFDYLDKGFGRQYLDLTRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 169
 QY 121 WGLASEILMOWNDAMEDLTKLKEIDNNSVSSPLOSQORTWLHWSLFFVFNHPKGR 180
 DB 170 WGLASEILMOWNDAMEDLTKLKEIDNNSVSSPLOSQORTWLHWSLFFVFNHPKGR 229
 QY 181 DNIIDLFLQPYLNAIQTCWCHILRYLTVTAVITNKDVKRQVLKDLVKVLCQESYTYK 240
 DB 230 DNIIDLFLQPYLNAIQTCWCHILRYLTVTAVITNKDVKRQVLKDLVKVLCQESYTYK 289
 QY 241 DPITEFVECLYVNFDFGAKKLRECESVLVNDFLVLACLEDIENARLFIPTFCRIHQ 300
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 290 DPITEFVECLYVNFDFGAKKLRECESVLVNDFLVLACLEDIENARLFIPTFCRIHQ 349
 QY 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIEKTK 360
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 350 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIEKTK 409
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 410 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 15
 ABU05157
 ID ABU05157 standard; protein; 445 AA.
 XX AC ABU05157;
 XX DT 29-JAN-2003 (first entry)
 XX DE Human expressed protein tag (EPT) #1823.
 XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX OS Homo sapiens.
 XX PN WO200278524-A2.
 XX PD 10-OCT-2002.
 XX PF 28-MAR-2002; 2002WO-US009671.
 XX PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX (ZYCO-) ZYCOS INC.
 XX Chicx RM, Tomlinson AJ, Urban RG;
 XX WPI; 2003-040607/03.
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX Example 2; SEQ ID NO 1823; 134pp; English.
 XX The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 445 AA;

Query Match 100.0%; Score 396; DB 6; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVD FAMD VYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMQSTRD 60
Db |||||
Qy 50 MVD FAMD VYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMQSTRD 109
Db |||||
Qy 61 GRMLFDYLADKHGFRQEVLDLTRYAKFYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
Db |||||
Qy 110 GRMLFDYLADKHGFRQEVLDLTRYAKFYECGNYSGAAEYLYFFRVLVPATDRNALSSL 169
Db |||||
Qy 121 WGLKASEILMONWDAAMEDLTRLKETIDNNSVSPLOSLOQRTWLIHWSLVFFNHPKGR 180
Db |||||
Qy 170 WGLKASEILMONWDAAMEDLTRLKETIDNNSVSPLOSLOQRTWLIHWSLVFFNHPKGR 229
Db |||||
Qy 181 DNIIDFLYQPYQLNAIQTCMCPHILRLVLTAVITNKDVRKRRQVLKDLVKVQQESYTYK 240
Db |||||
Qy 230 DNIIDFLYQPYQLNAIQTCMCPHILRLVLTAVITNKDVRKRRQVLKDLVKVQQESYTYK 289
Db |||||
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECSVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
Db |||||
Qy 290 DPITEFVECLYVNFDFDGAQKLRCECSVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 349
Db |||||
Qy 301 CISINMLADKLNMTPPEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Db |||||
Qy 350 CISINMLADKLNMTPPEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 409
Db |||||
Qy 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db |||||
Qy 410 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445
Db |||||

Search completed: March 8, 2005, 20:22:44
Job time : 168 secs


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:23:41 ; Search time 137 Seconds
(without alignments)
951.108 Million cell updates/sec

Title: US-10-783-415-4
Perfect score: 396
Sequence: 1 MVDPMVDVKNLYSDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table:  Gapop 60.0, Gapext 60.0

Searched: 1391452 seqs, 329044822 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	396	100.0	396	10	US-09-858-152A-2
2	396	100.0	396	10	US-09-858-152A-4
3	396	100.0	396	16	US-10-408-765A-2064
4	396	100.0	396	16	US-10-783-415-2
5	396	100.0	396	16	US-10-783-415-4
6	396	100.0	445	13	US-10-087-192-558
7	204	51.5	320	13	US-10-087-192-555
8	103	26.0	117	15	US-10-264-049-4238
9	45	11.4	45	9	US-09-864-761-42538
10	42	10.6	42	9	US-09-864-761-40729
11	42	10.6	42	9	US-09-864-761-42138
12	34	8.6	112	14	US-10-106-698-6920
13	32	3.0	437	15	US-10-424-599-266298

14	10	2.5	439	16	US-10-437-963-188307	Sequence 188307,
15	10	2.5	439	16	US-10-767-701-45301	Sequence 45301, A
16	10	2.5	444	16	US-10-437-963-188308	Sequence 188308,
17	9	2.3	9	14	US-10-022-066-57	Sequence 57, Appl
18	9	2.3	9	14	US-10-022-066-395	Sequence 395, App
19	9	2.3	130	15	US-10-424-599-217317	Sequence 217317,
20	9	2.3	147	15	US-10-425-114-49583	Sequence 49583, A
21	9	2.3	187	15	US-10-425-114-61814	Sequence 61814, A
22	8	2.0	401	15	US-10-424-599-243201	Sequence 243201, A
23	8	2.0	4092	15	US-10-369-493-22278	Sequence 22278, A
24	7	1.8	44	14	US-10-073-118-28	Sequence 28, Appl
25	7	1.8	45	14	US-10-289-135A-34	Sequence 34, Appl
26	7	1.8	57	15	US-10-424-599-192126	Sequence 192126,
27	7	1.8	70	14	US-10-029-386-30307	Sequence 30307, A
28	7	1.8	74	14	US-10-279-964-8	Sequence 8, Appl
29	7	1.8	75	15	US-10-424-599-240685	Sequence 240685,
30	7	1.8	81	15	US-10-291-172-276	Sequence 276, App
31	7	1.8	81	15	US-10-221-278-276	Sequence 276, App
32	7	1.8	81	15	US-10-468-125-12	Sequence 12, Appl
33	7	1.8	87	15	US-10-424-599-189674	Sequence 189674,
34	7	1.8	89	15	US-10-291-172-652	Sequence 652, App
35	7	1.8	89	15	US-10-221-278-652	Sequence 652, App
36	7	1.8	108	15	US-10-424-599-171619	Sequence 171619,
37	7	1.8	122	16	US-10-437-963-191022	Sequence 191022,
38	7	1.8	131	15	US-10-424-599-190181	Sequence 190181,
39	7	1.8	133	9	US-09-870-844-6	Sequence 6, Appl
40	7	1.8	146	16	US-10-437-963-192808	Sequence 192808,
41	7	1.8	179	14	US-10-316-253-295	Sequence 295, App
42	7	1.8	181	10	US-09-847-208-154	Sequence 154, App
43	7	1.8	181	14	US-10-316-253-297	Sequence 297, App
44	7	1.8	181	15	US-10-240-240A-10	Sequence 10, Appl
45	7	1.8	182	9	US-09-374-671-89	Sequence 89, Appl

ALIGNMENTS

RESULT 1
US-09-858-152A-2
; Sequence 2, Application US/09858152A
; Publication No. US2003004419A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE
; SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858.152A
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Murine INT6
US-09-858-152A-2

Query Match	100.0%	Score	396;	DB	10;	Length	396;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	396;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MVDPMVDVKNLYSDDIPHA	REKRTTVAAQKQQAETEP	IVKMFEDP	ETTRQMQSTRD	60	
Db	1	MVDPMVDVKNLYSDDIPHA	REKRTTVAAQKQQAETEP	IVKMFEDP	ETTRQMQSTRD	60	
Qy	61	GRMLFDYLDKGGFRQBYDL	TLRYAKFQYECG	NYGAAEYLYFRVL	VPATDNNALSSL	120	
Db	61	GRMLFDYLDKGGFRQBYDL	TLRYAKFQYECG	NYGAAEYLYFRVL	VPATDNNALSSL	120	
Qy	121	WGKLASEILMNWDAMEDL	TRLKETDNNNSV	SSPLSQLO	RTWLIIHWSL	VFVFNHPKGR	180

```
Db 121 WGKLASEILMKNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIHWSLFVFNHPKGR 180
Qy 181 DNIIDFLYQPOYLNAIQTCMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Db 181 DNIIDFLYQPOYLNAIQTCMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 2
US-09-858-152A-4
; Sequence 4, Application US/09858152A
; Publication No. US2003004419A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE
; APPLICANT: SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858,152A
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-152A-4

Query Match 100.0%; Score 396; DB 10; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMPEDPETTRQMOSTRD 60
Db 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMPEDPETTRQMOSTRD 60
Qy 61 GRMLFDYLADKHGFRQEYLDTRYAKFQYECGNYSGAAEYLYFFRVLPATDRLNALSSL 120
Db 61 GRMLFDYLADKHGFRQEYLDTRYAKFQYECGNYSGAAEYLYFFRVLPATDRLNALSSL 120
Qy 121 WGKLASEILMKNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIHWSLFVFNHPKGR 180
Db 121 WGKLASEILMKNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIHWSLFVFNHPKGR 180
Qy 181 DNIIDFLYQPOYLNAIQTCMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Db 181 DNIIDFLYQPOYLNAIQTCMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 3
US-10-408-765A-2064
; Sequence 2064, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2064
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2064

Query Match 100.0%; Score 396; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMPEDPETTRQMOSTRD 60
Db 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMPEDPETTRQMOSTRD 60
Qy 61 GRMLFDYLADKHGFRQEYLDTRYAKFQYECGNYSGAAEYLYFFRVLPATDRLNALSSL 120
Db 61 GRMLFDYLADKHGFRQEYLDTRYAKFQYECGNYSGAAEYLYFFRVLPATDRLNALSSL 120
Qy 121 WGKLASEILMKNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIHWSLFVFNHPKGR 180
Db 121 WGKLASEILMKNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIHWSLFVFNHPKGR 180
Qy 181 DNIIDFLYQPOYLNAIQTCMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Db 181 DNIIDFLYQPOYLNAIQTCMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 4
US-10-783-415-2
; Sequence 2, Application US/10783415
; Publication No. US20040141918A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
```

```
Db 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 3
US-10-408-765A-2064
; Sequence 2064, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2064
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2064

Query Match 100.0%; Score 396; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMPEDPETTRQMOSTRD 60
Db 1 MVDFAVDVYKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMPEDPETTRQMOSTRD 60
Qy 61 GRMLFDYLADKHGFRQEYLDTRYAKFQYECGNYSGAAEYLYFFRVLPATDRLNALSSL 120
Db 61 GRMLFDYLADKHGFRQEYLDTRYAKFQYECGNYSGAAEYLYFFRVLPATDRLNALSSL 120
Qy 121 WGKLASEILMKNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIHWSLFVFNHPKGR 180
Db 121 WGKLASEILMKNQWDAAMEDLTRKETIDNNSVSPLOSQORTWLIHWSLFVFNHPKGR 180
Qy 181 DNIIDFLYQPOYLNAIQTCMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Db 181 DNIIDFLYQPOYLNAIQTCMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 4
US-10-783-415-2
; Sequence 2, Application US/10783415
; Publication No. US20040141918A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
```

; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/10/783,415
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 09/858,152
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Murine INT6
US-10-783-415-2

Query Match 100.0%; Score 396; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRQMQSTRD 60
DB 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRQMQSTRD 60

QY 61 GRMLFDYLADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
DB 61 GRMLFDYLADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120

QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 180
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 180

QY 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTYK 240
DB 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTYK 240

QY 241 DPITEFVECLYVNFDFDGAQKLRCEESLVNDPFLVACLEDFIENARLFIFETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESLVNDPFLVACLEDFIENARLFIFETFCRIHQ 300

QY 301 CISINMLADKLNMTPEEAERWIVNIRNARDAKIDSKLGHVVMGNNNAVSPYQOVIETK 360
DB 301 CISINMLADKLNMTPEEAERWIVNIRNARDAKIDSKLGHVVMGNNNAVSPYQOVIETK 360

QY 361 SLSFRSQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
DB 361 SLSFRSQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 5

US-10-783-415-4
; Sequence 4, Application US/10783415
; Publication No. US20040141918A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/10/783,415
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 09/858,152
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-783-415-4

Query Match 100.0%; Score 396; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRQMQSTRD 60
DB 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRQMQSTRD 60

QY 61 GRMLFDYLADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
DB 61 GRMLFDYLADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120

QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 180
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 180

QY 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTYK 240
DB 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDVKRRQVLDLVKVIQOESYTYK 240

QY 241 DPITEFVECLYVNFDFDGAQKLRCEESLVNDPFLVACLEDFIENARLFIFETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESLVNDPFLVACLEDFIENARLFIFETFCRIHQ 300

QY 301 CISINMLADKLNMTPEEAERWIVNIRNARDAKIDSKLGHVVMGNNNAVSPYQOVIETK 360
DB 301 CISINMLADKLNMTPEEAERWIVNIRNARDAKIDSKLGHVVMGNNNAVSPYQOVIETK 360

QY 361 SLSFRSQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
DB 361 SLSFRSQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 6

US-10-087-192-558
; Sequence 558, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 558
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-558

Query Match 100.0%; Score 396; DB 13; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRQMQSTRD 60
DB 50 MVDFAVDYKLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFDPETTRQMQSTRD 109

QY 61 GRMLFDYLADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
DB 110 GRMLFDYLADKHGFRQEVLDLTRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 169

QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMQNWDAAEMDLTRLKETIDNNSSVSPQLSQLOQTWLIHWSLFFVFNHPKGR 229

Qy 181 DNIIDFLYQYQYLNAIQTMCPHILRYLTAVITNKQVRKROVLKOLVKVIOQESYTYK 240
Db 230 DNIIDFLYQYQYLNAIQTMCPHILRYLTAVITNKQVRKROVLKOLVKVIOQESYTYK 289
Qy 241 DPITEFVECLVNFDFDGAQKLRCECVLVNDFLVACLEDPIENARLFIFETFCRIHQ 300
Db 290 DPITEFVECLVNFDFDGAQKLRCECVLVNDFLVACLEDPIENARLFIFETFCRIHQ 349
Qy 301 CISINMLADKLNMTPEEARWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 360
Db 350 CISINMLADKLNMTPEEARWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 409
Qy 361 SLSFRSOMLANNIEKLNQNSRSEAPNWTQDSGFY 396
Db 410 SLSFRSOMLANNIEKLNQNSRSEAPNWTQDSGFY 445

RESULT 7
US-10-087-192-555
; Sequence 555, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 555
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-555

Query Match 51.5%; Score 204; DB 13; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.8e-186;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 193 YLNAIQTMCPHILRYLTAVITNKQVRKROVLKOLVKVIOQESYTYKDPITEFVECLYV 252
Db 117 YLNAIQTMCPHILRYLTAVITNKQVRKROVLKOLVKVIOQESYTYKDPITEFVECLYV 176
Qy 253 NFDFDGAQKLRCECVLVNDFLVACLEDPIENARLFIFETFCRIHQCSINNMLADKLN 312
Db 177 NFDFDGAQKLRCECVLVNDFLVACLEDPIENARLFIFETFCRIHQCSINNMLADKLN 236
Qy 313 MTPPEARWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTKSLSFRSOMLANN 372
Db 237 MTPPEARWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTKSLSFRSOMLANN 296
Qy 373 IEKLNQNSRSEAPNWTQDSGFY 396
Db 297 IEKLNQNSRSEAPNWTQDSGFY 320

RESULT 8
US-10-264-049-4238
; Sequence 4238, Application US/10264049
; Publication No. US20040000579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P413P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569

; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4238
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4238
Query Match 26.0%; Score 103; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 4e-90;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 126 SEILMQNWDAAEDLTRELKETIDNNSSVSPLOSQQRTWLHWSLFFVFNHPKGRDNIID 185
Db 5 SEILMQNWDAAEDLTRELKETIDNNSSVSPLOSQQRTWLHWSLFFVFNHPKGRDNIID 64
Qy 186 LFLYQPOYLNAIQTMCPHILRYLTAVITNKQVRKROVLKDL 228
Db 65 LFLYQPOYLNAIQTMCPHILRYLTAVITNKQVRKROVLKDL 107
RESULT 9
US-09-864-761-42538
; Sequence 42538, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42538
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP001331.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 24
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.8
; OTHER INFORMATION: EST_HUMAN HIT: A1718630.1, EVALUE 3.00e-18
; OTHER INFORMATION: SWISSPROT HIT: Q64252, EVALUE 2.00e-19
US-09-864-761-42538

Query Match 11.4%; Score 45; DB 9; Length 45;
Best Local Similarity 100.0%; Pred. No. 4.7e-35; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 GHVVGNNVSPYQOVIETKSLSPRSQMLAMNIEKLNQNSRSE 45
|||||

RESULT 10
US-09-864-761-40729
; Sequence 40729, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40729
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC020991.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EST_HUMAN HIT: BF700473.1, EVALUE 7.00e-17
; OTHER INFORMATION: SWISSPROT HIT: Q64252, EVALUE 6.00e-18
US-09-864-761-40729

Query Match 10.6%; Score 42; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 3.3e-32; Indels 0; Gaps 0;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 YLNAIQTMCPHILRYLTAVITNKDVKRRQVLKDLVKVIOQ 42
|||||

RESULT 11
US-09-864-761-42138
; Sequence 42138, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30

* INVENTOR: G. Msz
* APPLICANT: Bookharov, Andrey A.
* APPLICANT: Barbazuk, Brad
* APPLICANT: Li, Ping
* TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
* TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
* FILE REFERENCE: 38-21(53221)B
* CURRENT APPLICATION NUMBER: US/10/437.963

#

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188307
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84924C.1.pep
US-10-437-963-188307

Query Match 2.5%; Score 10; DB 16; Length 439;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 285 ENARLFIFET 294
Db 339 ENARLFIFET 348

RESULT 15
US-10-767-701-45301
; Sequence 45301, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45301
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C1020_1.pep
US-10-767-701-45301

Query Match 2.5%; Score 10; DB 16; Length 439;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 285 ENARLFIFET 294
Db 339 ENARLFIFET 348

Search completed: March 8, 2005, 20:38:23
Job time : 138 secs

THIS PAGE BLANK (USSTO)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 19:47:46 ; Search time 172 Seconds

(without alignments)
1178.973 Million cell updates/sec

Title: US-10-783-415-4

Perfect score: 2071

Sequence: 1 MVDFAMDVYKNLYSDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_prot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2071	100.0	396	Q9BRV2	Q9brv2 homo sapien
2	2071	100.0	445	1 IF36 HUMAN	P60228 homo sapien
3	2071	100.0	445	1 IF36_MOUSE	P60229 mus musculus
4	2071	100.0	445	2 Q641X8	Q641x8 rattus norv
5	2067	99.8	445	2 Q8WVK4	Q8wvk4 homo sapien
6	2064	99.7	445	2 Q61AX5	Q61ax5 homo sapien
7	2006.5	96.9	446	2 Q6P7L9	Q6p7l9 xenopus tro
8	1992.5	96.2	423	2 Q7ZXA5	Q7zxa5 xenopus lae
9	1981.5	95.7	446	2 Q918W7	Q918w7 xenopus lae
10	1952.5	94.3	446	2 Q6P7X8	Q6p7x8 brachydanio
11	1947.5	94.0	446	2 Q6DR11	Q6dr11 brachydanio
12	1946	94.0	422	2 Q9CT23	Q9ct23 mus musculus
13	1783.5	86.1	448	2 Q6TH16	Q6th16 brachydanio
14	1220.5	58.9	435	1 IF36_DROME	Q77410 drosophila
15	1165.5	56.3	462	2 Q7Q1L8	Q7q1l8 anopheles g
16	1081.5	52.2	439	2 Q8GV10	Q8gv10 oryza sativ
17	1066	51.5	372	2 Q8MR88	Q8mr88 drosophila
18	1049.5	50.7	441	2 Q9C5Z3	Q9c5z3 arabidopsis
19	1048.5	50.6	441	2 Q9W4T7	Q9w4t7 arabidopsis
20	1011.5	48.8	415	2 Q69W36	Q69w36 oryza sativ
21	994.5	48.0	432	1 IF36_CAEEL	Q61820 caenorhabdi
22	990.5	47.8	397	2 Q69W37	Q69w37 oryza sativ
23	927.5	44.8	436	2 Q62475	Q62475 oryza sativ
24	903	43.6	418	2 Q9M2L8	Q9m2l8 arabidopsis
25	791.5	38.2	443	2 Q7S519	Q7s519 neurospora
26	781	37.7	421	2 Q6C0Q2	Q6c0q2 yarrowia li
27	775	37.4	501	1 IF36_SCHPO	Q94513 schizosacch
28	553	26.7	119	2 Q8BN6E	Q8bne6 mus musculus
29	424	20.5	527	2 Q7RQ22	Q7rq22 plasmodium
30	413	19.9	517	2 Q813I5	Q813i5 plasmodium
31	396.5	19.1	405	2 Q9NEE7	Q9nee7 leishmania

32 377 18.2 73 2 Q6LCK2
33 317 15.3 165 2 Q6XIT2
34 136 6.6 817 2 Q82Y53
35 133 6.4 467 2 Q7RLA3
36 130.5 6.3 414 2 Q6BWI2
37 123 5.9 604 2 Q6C1F1
38 118.5 5.7 2366 2 Q7RHV9
39 116.5 5.6 437 1 CSN2_SCHPO
40 116.5 5.6 7500 2 Q7KTP3
41 116.5 5.6 7735 2 Q9VMT8
42 116 5.6 458 1 GYP6_YEAST
43 116 5.6 654 2 Q6D8B3
44 115.5 5.6 482 2 Q8WRS9
45 115 5.6 444 1 CSN2_DROME

ALIGNMENTS

RESULT 1
Q9BRV2

ID Q9BRV2 PRELIMINARY; PRT; 396 AA.

AC Q9BRV2; 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE E1F3S6 protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.243603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RT Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences."

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RA Strausberg R.;

RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC005944; AA05944.1; -.

DR InterPro; IPR000717; PCI.

DR InterPro; IPR010935; SMC_hinge.

DR Pfam; PF01399; PCI; 1.

DR SMART; SM00088; PINT; 1.

SQ SEQUENCE 396 AA; 46445 MW; 09165D0D612B1C3B CRC64;

Query Match 100.0%; Score 2071; DB 2; Length 396;

Best Local Similarity 100.0%; Pred. No. 5e-145;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAMDVYKNLYSDDIPHA.....LNQNSRSEAPNWTQDSGFYVVAQKQIQAEETPIVKPFEDPTTROMOSTRD 60

Db 1 MVDAMDVYKNLYSDSDIPHALREKRTTVVAAQLKQIQAETPIVKMFEDPETTRQWQSTRD 60
Qy 61 GRMLFDYLDKRGFRQEVLDLYRYAKFQYCGNYSGAAYLYYFRVLVLPATDRNALSSL 120
Db 61 GRMLFDYLDKRGFRQEVLDLYRYAKFQYCGNYSGAAYLYYFRVLVLPATDRNALSSL 120
Qy 121 WGLKASEITLQWDAAMEDTLRLKETIDNNSSVSSPLOSLOQRTWLHWSLFFVFNHPKGR 180
Db 121 WGLKASEITLQWDAAMEDTLRLKETIDNNSSVSSPLOSLOQRTWLHWSLFFVFNHPKGR 180
Qy 181 DNIIDLFLYQPOLNATQTMCPHILRYLTAVITNKDVRKQVLKDLVKVVIQESYTYK 240
Db 181 DNIIDLFLYQPOLNATQTMCPHILRYLTAVITNKDVRKQVLKDLVKVVIQESYTYK 240
Qy 241 DPTEFVECLYVNFDFDGAQKLRCESEVNVNDFLVAACLEDFTENARLFTFETPCRHHQ 300
Db 241 DPTEFVECLYVNFDFDGAQKLRCESEVNVNDFLVAACLEDFTENARLFTFETPCRHHQ 300
Qy 301 CISTINMLADKLNMTPEERWIVNLRNARDLDAKIDSLGHVVMGNNVSPVQQVIEKTK 360
Db 301 CISTINMLADKLNMTPEERWIVNLRNARDLDAKIDSLGHVVMGNNVSPVQQVIEKTK 360
Qy 361 SLSPRSQMLAMNIEKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSPRSQMLAMNIEKLNQNSRSEAPNATQDSGFY 396

RESULT 2
ID IF36 HUMAN STANDARD; PRT: 445 AA.
AC P60228; O43902; Q64058; Q64059; Q64252;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48)
DE (eIF3e) (Viral integration site protein INT-6 homolog).
GN Name=EIF3S6; Synonym=INT6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 280-289 AND 427-436.
RX TISSUE=Liver;
RX MEDLINE=97442403; PubMed=9295280; DOI=10.1074/jbc.272.38.23477;
RA Asano K., Merrick W.C., Hershey J.W.B.;
RT "The translation initiation factor eIF3-p48 subunit is encoded by int-6, a site of frequent integration by the mouse mammary tumor virus Genome.";
RL J. Biol. Chem. 272:23477-23480(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung;
RX MEDLINE=98066777; PubMed=9403073; DOI=10.1006/geno.1997.4996;
RA Miyazaki S., Imatani A., Ballard L., Marchetti A., Buttitta F.,
RA Albertsen H., Nevanlinna H.A., Gallahan D., Callahan R.;
RT "The chromosome location of the human homolog of the mouse mammary tumor-associated gene INT6 and its status in human breast carcinomas.";
RL Genomics 46:155-158(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96337990; PubMed=8688078;
RA Desbois C., Rousset R., Bantignies F., Jalinot P.;
RT "Exclusion of Int-6 from PML nuclear bodies by binding to the HTLV-I Tax oncoprotein.";
RL Science 273:951-953(1996).
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=12386384;
RA Neuvert C., Jin D.-Y., Semmes O.J., Diella F., Callahan R.,
RA Jeang K.-T.;
RT "Divergent subcellular locations of HTLV-I Tax and Int-6: a contrast

RT between in vitro protein-protein binding and intracellular protein colocalization.";
RL J. Biomed. Sci. 4:229-234(1997).
RN [5]
RP SEQUENCE FROM N.A.
RX TISSUE=Bone marrow, Brain, and Muscle;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.J., Max S.I., Wang J., Hsieh F.,
RA Stapchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Buttrick A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP INTERACTION WITH TRIM27.
RX PubMed=10504338;
RA Morris-Desbois C., Bochar V., Reynaud C., Jalinot P.;
RT "Interaction between the Ret finger protein and the Int-6 gene product and co-localisation into nuclear bodies.";
RL Cell Sci. 112:3331-3342(1999).
CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of methionyl-tRNAi and mRNA.
CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits.
CC -!- INTERACTS WITH TRIM27.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- SIMILARITY: Belongs to the EIF3S6 family.
CC -!- SIMILARITY: Contains 1 PCI domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U54562; AAC51760.1; -;
CC EMBL; U94174; AAC51917.1; -;
CC EMBL; U94162; AAC51917.1; JOINED.
CC EMBL; U94163; AAC51917.1; JOINED.
CC EMBL; U94164; AAC51917.1; JOINED.
CC EMBL; U94165; AAC51917.1; JOINED.
CC EMBL; U94166; AAC51917.1; JOINED.
CC EMBL; U94167; AAC51917.1; JOINED.
CC EMBL; U94168; AAC51917.1; JOINED.
CC EMBL; U94169; AAC51917.1; JOINED.
CC EMBL; U94170; AAC51917.1; JOINED.
CC EMBL; U94171; AAC51917.1; JOINED.
CC EMBL; U94172; AAC51917.1; JOINED.
CC EMBL; U94173; AAC51917.1; JOINED.
CC EMBL; U94175; AAC51919.1; -;
CC EMBL; U62962; AAB58251.1; -;
CC EMBL; U85947; AAB88873.1; -;
CC EMBL; BC000734; AAH00734.1; -;
CC EMBL; BC008419; AAH08419.1; -;
CC EMBL; BC016706; AAH16706.1; -;
CC EMBL; BC021679; AAH21679.1; -;
CC IntAct; P60228; -;

DR Genew; HGNC:3277; EIF3S6.
DR H-InVDB; HIX0007722; --
DR Reactome; P60228; --
DR MIM; 602210; --
DR InterPro; IPR000717; PCI.
KW Direct protein sequencing; Initiation factor; Protein biosynthesis.
SQ SEQUENCE 445 AA; 5220 MW; A5368651DDDDDD0C CRC64;

Query Match 100.0%; Score 2071; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 5.8e-145;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPALREKRTTVAQLKQLQAEPIVYKMFEDPETTRQMQSTRD 60
DB 50 MVDFAVDVYKNLYSDDIIPALREKRTTVAQLKQLQAEPIVYKMFEDPETTRQMQSTRD 109
QY 61 GRMLFDYLDKXGFRQYEDLTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAIDRNALSSL 120
DB 110 GRMLFDYLDKXGFRQYEDLTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAIDRNALSSL 169
QY 121 WGLKASEILMKNWDAAMEDLTRKETIDNNSVSSPLOSQOQRTWLIHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMKNWDAAMEDLTRKETIDNNSVSSPLOSQOQRTWLIHWSLFFVFNHPKGR 229
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRQVLKDLVKVYQESYTYK 240
DB 230 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRQVLKDLVKVYQESYTYK 289
QY 241 DPTFVECLYVNFDPGAKKLRECVSLVNDFFLVACLEDPIENARLFIPTFCRIHQ 300
DB 290 DPTFVECLYVNFDPGAKKLRECVSLVNDFFLVACLEDPIENARLFIPTFCRIHQ 349
QY 301 CTSINMLADKLNTPEASRWIINLRNARLDKIDSKLGHVYMGNNVSPQVVIETK 360
DB 350 CTSINMLADKLNTPEASRWIINLRNARLDKIDSKLGHVYMGNNVSPQVVIETK 409
QY 361 SLSFRSOMLAMIETKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSOMLAMIETKLNQNSRSEAPNWTQDSGFY 445

RESULT 3
ID IF36 MOUSE STANDARD; PRT; 445 AA.
AC P60229; Q64058; Q64059; Q64252;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48)
DE (eIF3e) (Mammary tumor-associated protein INT-6) (Viral integration site protein INT-6) (MMTV integration site 6);
GN Name=EIF3e; Synonyms=Int6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95156630; PubMed=7853537;
RA Marchetti A., Buttitta F., Miyazaki S., Callahan D., Smith G.H., Callahan R.;
RT "Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia.";
RL J. Virol. 69:1932-1938 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97405883; PubMed=9260927;
RA Diella F., Levi G., Callahan R.;
RT "Characterization of the INT6 mammary tumor gene product.";
RL DNA Cell Biol. 16:839-847 (1997).
RN [3]
RP REVISIONS TO N-TERMINUS.
RA Callahan R.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=Czech II; TISSUE=Mammary gland;
RC MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M.J., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalish D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of methionyl-tRNAi and mRNA.
CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits. Interacts with TRIM27 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- DISEASE: Int-6 serves as a site for viral integration of mouse mammary tumor virus (MMTV) in mammary tumors.
CC -!- SIMILARITY: Belongs to the EIF3E family.
CC -!- SIMILARITY: Contains 1 PCI domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S75221; AAC00046.1; --
CC EMBL; S75223; AAC00047.1; --
DR EMBL; BC029177; AAH29177.1; --
DR MGD; MGI:99257; Bif3a6.
DR InterPro; IPR000717; PCI.
KW Initiation factor; Protein biosynthesis; Proto-oncogene.
FT VARIANT 158 158
FT VARIANT 158 180
FT VARIANT 159 445
FT VARIANT 318 365
FT VARIANT 366 445
FT VARIANT 445 AA; 52220 MW; A5368651DDDDDD0C CRC64;
SQ SEQUENCE 445 AA; 52220 MW; A5368651DDDDDD0C CRC64;

Query Match 100.0%; Score 2071; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 5.8e-145;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPALREKRTTVAQLKQLQAEPIVYKMFEDPETTRQMQSTRD 60
DB 50 MVDFAVDVYKNLYSDDIIPALREKRTTVAQLKQLQAEPIVYKMFEDPETTRQMQSTRD 109
QY 61 GRMLFDYLDKXGFRQYEDLTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAIDRNALSSL 120
DB 110 GRMLFDYLDKXGFRQYEDLTLYRYAKFQYECGNYSGAAEYLYFFRVLVPAIDRNALSSL 169

```

QY 121 WGLASEILMNQWDAAMEDTLRLKETIDNNSSVSSPLQSQLOORTWLIIHWSLFFVFNHPKGR 180
DB 170 WGLASEILMNQWDAAMEDTLRLKETIDNNSSVSSPLQSQLOORTWLIIHWSLFFVFNHPKGR 229
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQSSYTYK 240
DB 230 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQSSYTYK 289
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 349
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEEAERWIVNLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSQMLAMNIEKLNQNSRSEAPNWTODSGFY 396
DB 410 SLSFRSQMLAMNIEKLNQNSRSEAPNWTODSGFY 445

RESULT 4
Q641X8
ID Q641X8 PRELIMINARY; PRT; 445 AA.
AC Q641X8;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082087; AAH82087.1; -.
KW Hypothetical protein.
SQ SEQUENCE 445 AA; 52220 MW; A5368651DDDDDDOC CRC64;

Query Match 100.0%; Score 2071; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 5.8e-145;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDPMDDVYKNLYSDDIPHALREKRTTVVAQLKQLAQATEPIVKMFPEDPTTRQMOSTRD 60
DB 50 MVDPMDDVYKNLYSDDIPHALREKRTTVVAQLKQLAQATEPIVKMFPEDPTTRQMOSTRD 109

```

```

QY 61 GRMLFDYLDADKHGFRQEYLDLTRYAKFQYECGNYSGAAEYLYFFRVLPATRNALSSL 120
DB 110 GRMLFDYLDADKHGFRQEYLDLTRYAKFQYECGNYSGAAEYLYFFRVLPATRNALSSL 169
QY 121 WGLASEILMNQWDAAMEDTLRLKETIDNNSSVSSPLQSQLOORTWLIIHWSLFFVFNHPKGR 180
DB 170 WGLASEILMNQWDAAMEDTLRLKETIDNNSSVSSPLQSQLOORTWLIIHWSLFFVFNHPKGR 229
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQSSYTYK 240
DB 230 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQSSYTYK 289
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 349
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEEAERWIVNLRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSQMLAMNIEKLNQNSRSEAPNWTODSGFY 396
DB 410 SLSFRSQMLAMNIEKLNQNSRSEAPNWTODSGFY 445

RESULT 5
Q6WVK4
ID Q6WVK4 PRELIMINARY; PRT; 445 AA.
AC Q6WVK4;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Mammmary tumor integration site 6 (Oncogene homolog).
GN Name=E1F3S6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017887; AAH17887.1; -.
DR PFam; PF01399; PCl; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 445 AA; 52248 MW; AF0AD651DAAA8D0C CRC64;

```



```

Query Match          99.8%; Score 2067; DB 2; Length 445;
Best Local Similarity 99.7%; Pred. No. 1.1e-144;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MYDFAMDVYKNLYSDDI PHALREKRTTVAQKQLQAEETPIVKMFEDPETTRQMOSTRD 60
DB 50 MYDFAMDVYKNLYSDDI PHALREKRTTVAQKQLQAEETPIVKMFEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKGGFQREYLDLTRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
DB 110 GRMLFDYLDKGGFQREYLDLTRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 169
QY 121 WGLKASEILMQNWDAAEDLTRKETIDNNSVSSPLQSIQORTWLIHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMQNWDAAEDLTRKETIDNNSVSSPLQSIQORTWLIHWSLFFVFNHPKGR 229
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRROVLKDLVKVIOQESYTYK 240
DB 230 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRROVLKDLVKVIOQESYTYK 289
QY 241 DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 349
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETKT 360
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETKT 409
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 6
Q6IAX5 PRELIMINARY; PRT; 445 AA.
AC Q6IAX5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Eukaryotic translation initiation factor 3, subunit 6.
DE EIF3S6 protein.
GN Name=EIF3S6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR457029; CNG33310.1; -.
DR InterPro; IPR000717; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 445 AA; 52205 MW; C7C0FF75BD9A8D4 CRC64;

Query Match          99.7%; Score 2064; DB 2; Length 445;
Best Local Similarity 99.7%; Pred. No. 1.9e-144;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MYDFAMDVYKNLYSDDI PHALREKRTTVAQKQLQAEETPIVKMFEDPETTRQMOSTRD 60
DB 50 MYDFAMDVYKNLYSDDI PHALREKRTTVAQKQLQAEETPIVKMFEDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKGGFQREYLDLTRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
DB 110 GRMLFDYLDKGGFQREYLDLTRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 169
QY 121 WGLKASEILMQNWDAAEDLTRKETIDNNSVSSPLQSIQORTWLIHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMQNWDAAEDLTRKETIDNNSVSSPLQSIQORTWLIHWSLFFVFNHPKGR 229

```

```

QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRROVLKDLVKVIOQESYTYK 240
DB 230 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRROVLKDLVKVIOQESYTYK 289
QY 241 DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 349
QY 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETKT 360
DB 350 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETKT 409
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
DB 410 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 7
Q6P7L9 PRELIMINARY; PRT; 446 AA.
AC Q6P7L9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Eukaryotic translation initiation factor 3, subunit 6.
DE Name=elf3-p48-prov;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OX Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.H., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; SC061611; AAH61611.1; -.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
DR Initiation factor.
SQ SEQUENCE 446 AA; 52296 MW; EB12CF70898DA8B1 CRC64;

Query Match          96.9%; Score 2006.5; DB 2; Length 446;
Best Local Similarity 96.5%; Pred. No. 3.4e-140;
Matches 383; Conservative 9; Mismatches 4; Indels 1; Gaps 1;

```

```

QY 1 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMSTR 59
DB 50 MVDFAVDVYKNLYADKEIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMSTR 109
QY 60 DGRMLFDYLDADKHGFRQBYDLTYRYAKFQYECGNSGAAEYLYFFRVLVVPATDRNALSS 119
DB 110 DGRMLFDHLAEKHGFRQBYDLTYRYAKFQYECGNSGAAEYLYFFRVLVVPSTDRNALSS 169
QY 120 LGWKLASEILMQLNWDAAEMEDTLRLKETIDNNSSVPSLOQRTWLIIHWSLFFVFNHPKG 179
DB 170 LGWKLASEILMQLNWDAAEMEDTLRLKETIDNNSSVPSLOQRTWLIIHWSLFFVFNHPKG 229
QY 180 RDNIDFLYQOYLNAIQTCMCPHLRLYLTAVITNKDVRKROVLKDLVKVIOQESYTY 239
DB 230 RDNIDFLYQOYLNAIQTCMCPHLRLYLTAVITNKDVRKROVLKDLVKVIOQESYTY 289
QY 240 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFTIENARLFIPTFCRIH 299
DB 290 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFTIENARLFIPTFCRIH 349
QY 300 QCISINMLADKLNMTPPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 359
DB 350 QCISINMLADKLNMTPPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 409
QY 360 KSLSPRSQWLAMNIEKLNQNSRSEAPNWAOTDSGFY 396
DB 410 KSLAFRSQWLAMNIEKLNQNSRSEAPNWAOTDSGFY 446

RESULT 8
Q7ZXAS PRELIMINARY; PRT; 423 AA.
AC Q7ZXAS;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-WAR-2004 (TREMBlrel. 26, Last annotation update)
DE LOC398503 protein (Fragment).
GN Name=LOC398503;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson F.;

```

```

RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RT Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045079; AAH45079.1; -.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
FT NON TER 1
SQ SEQUENCE 423 AA; 49523 MW; F691314F43EF400B CRC64;

Query Match 96.2%; Score 1992.5; DB 2; Length 423;
Best Local Similarity 96.0%; Pred. No. 3.5e-139;
Matches 381; Conservative 10; Mismatches 5; Indels 1; Gaps 1;

QY 1 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMSTR 59
DB 27 MVDFAVDVYKNLYADKEIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMSTR 86
QY 60 DGRMLFDYLDADKHGFRQBYDLTYRYAKFQYECGNSGAAEYLYFFRVLVVPATDRNALSS 119
DB 87 DGRMLFDHLAEKHGFRQBYDLTYRYAKFQYECGNSGAAEYLYFFRVLVVPSTDRNALSS 146
QY 120 LGWKLASEILMQLNWDAAEMEDTLRLKETIDNNSSVPSLOQRTWLIIHWSLFFVFNHPKG 179
DB 147 LGWKLASEILMQLNWDAAEMEDTLRLKETIDNNSSVPSLOQRTWLIIHWSLFFVFNHPKG 206
QY 180 RDNIDFLYQOYLNAIQTCMCPHLRLYLTAVITNKDVRKROVLKDLVKVIOQESYTY 239
DB 207 RDNIDFLYQOYLNAIQTCMCPHLRLYLTAVITNKDVRKROVLKDLVKVIOQESYTY 266
QY 240 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFTIENARLFIPTFCRIH 299
DB 267 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFTIENARLFIPTFCRIH 326
QY 300 QCISINMLADKLNMTPPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 359
DB 327 QCISINMLADKLNMTPPEAEERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 386
QY 360 KSLSPRSQWLAMNIEKLNQNSRSEAPNWAOTDSGFY 396
DB 387 KSLAFRSQWLAMNIEKLNQNSRSEAPNWAOTDSGFY 423

RESULT 9
Q918W7 PRELIMINARY; PRT; 446 AA.
ID Q918W7 .. PRELIMINARY; PRT; 446 AA.
AC Q918W7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-WAR-2004 (TREMBlrel. 26, Last annotation update)
DE Int-6 protein. (African clawed frog).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Morris-Desbois C., Bochar V., Reynaud C., Jalinot P.;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162775; AAF80474.1; -.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 446 AA; 52240 MW; 21C5605DA9CF88BF CRC64;

```

```

Query Match      95.7%; Score 1981.5; DB 2; Length 446;
Best Local Similarity 95.2%; Pred. No. 2.4e-138;
Matches 378; Conservative 12; Mismatches 6; Indels 1; Gaps 1;

QY 1 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQKQLQAETEPVVKMFEDPETTRQMSTR 59
DB 50 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQKQLQAETEPVVKMFEDPETTRQMSTR 109

QY 60 DGRMLFDYLDKRGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSS 119
DB 110 DGRMLFDYLDKRGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSS 169

QY 120 LKGLKASEILMNQWDAAMEDLTRKETIDNNSSVSPLOSLQOQTWLIHWSLFFVFNHPKG 179
DB 170 LKGLKASEILMNQWDAAMEDLTRKETIDNNSSVSPLOSLQOQTWLIHWSLFFVFNHPKG 229

QY 180 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLVKVIQOESYTY 239
DB 230 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLVKVIQOESYTY 289

QY 240 KDPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIH 299
DB 290 KDPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIH 349

QY 300 QCSISINMLADKLNMTPEEAERWIVNLRNARLDKIDS KLGHVVMGNNAVSPYQQVIEKT 359
DB 350 QCSISINMLADKLNMTPEEAERWIVNLRNARLDKIDS KLGHVVMGNNAVSPYQQVIEKT 409

QY 360 KSLSFQSMQMLANNIEKKLNQNSRSEAPNATQDSGFY 396
DB 410 KSLSFQSMQMLANNIEKKLNQNSRSEAPNATQDSGFY 446

RESULT 10
Q6PX8 PRELIMINARY; PRT; 446 AA.
AC Q6PX8;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 3, subunit 6.
GN ORFNames=zgc:63821;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalak U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061454; AAH61454.1; -.
DR ZFIN; ZDB-GENE-030131-3827; zgc:63821.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
KW Initiation factor.
SQ SEQUENCE 446 AA; 52389 MW; 32F9D000E69F4F87 CRC64;

Query Match      94.3%; Score 1952.5; DB 2; Length 446;
Best Local Similarity 93.2%; Pred. No. 3.4e-136;
Matches 370; Conservative 18; Mismatches 8; Indels 1; Gaps 1;

QY 1 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQKQLQAETEPVVKMFEDPETTRQMSTR 59
DB 50 MVDFAVDVYKNLYSD-DIPHALREKRTTVVAQKQLQAETEPVVKMFEDPETTRQMSTR 109

QY 60 DGRMLFDYLDKRGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSS 119
DB 110 DGRMLFDYLDKRGFRQYELDTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSS 169

QY 120 LKGLKASEILMNQWDAAMEDLTRKETIDNNSSVSPLOSLQOQTWLIHWSLFFVFNHPKG 179
DB 170 LKGLKASEILMNQWDAAMEDLTRKETIDNNSSVSPLOSLQOQTWLIHWSLFFVFNHPKG 229

QY 180 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLVKVIQOESYTY 239
DB 230 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLVKVIQOESYTY 289

QY 240 KDPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIH 299
DB 290 KDPITEFVECLYVNFDFDGAOKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIH 349

QY 300 QCSISINMLADKLNMTPEEAERWIVNLRNARLDKIDS KLGHVVMGNNAVSPYQQVIEKT 359
DB 350 QCSISINMLADKLNMTPEEAERWIVNLRNARLDKIDS KLGHVVMGNNAVSPYQQVIEKT 409

QY 360 KSLSFQSMQMLANNIEKKLNQNSRSEAPNATQDSGFY 396
DB 410 KSLSFQSMQMLANNIEKKLNQNSRSEAPNATQDSGFY 446

RESULT 11
Q6DRI1 PRELIMINARY; PRT; 446 AA.
AC Q6DRI1;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Eukaryotic translation initiation factor 3 subunit 6.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15256591; DOI=10.1073/pnas.0403929101;
RA Amsterdam A., Nissen R.M., Sun Z., Swindell E.C., Farrington S.,
RA Hopkins N.;
RT "Identification of 315 genes essential for early zebrafish
RT development."
RL Proc. Natl. Acad. Sci. U.S.A. 101:12792-12797(2004).
DR EMBL; AF648778; AAF68096.1; -.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.

```

```
KW Initiation factor.
SQ SEQUENCE 446 AA; 52398 MW; 2C36E46A8002BB7 CRC64;

Query Match 94.0%; Score 1947.5; DB 2; Length 446;
Best Local Similarity 92.9%; Pred. No. 8e-136;
Matches 369; Conservative 18; Mismatches 9; Indels 1; Gaps 1;

QY 1 MYDFAMDVYKNLYSD-DIPHALREKRTTVVAOLKQLOAETEPVVKMFEDPETTRQMOSTR 59
DB 50 MYDFAMDVYKNLYSD-DIPHALREKRTTVVAOLKQLOAETEPVVKMFEDPETTRQMOSTR 109

QY 60 DGRMLFDYLADKHGRFQRYELDTLYRYAKFQYECGNYSGAAEYLYFRVLVPAVDNRNALSS 119
DB 110 DGRMLFDYLADKHGRFQRYELDTLYRYAKFQYECGNYSGAAEYLYFRVLVPAVDNRNALSS 169

QY 120 LMGKLAISEILMQNWDAAEMDLTRLKETIDNNVSSPLQSLQORTWLIHWSLFFVFNHPKG 179
DB 170 LMGKLAISEILMQNWEAEMDLTRLKETIDNNVSSPLQSLQORTWLIHWSLFFVFNHPKG 229

QY 180 RNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKQVRKROVLKDLVKVIOQESYTY 239
DB 230 RNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKQVRKROVLKDLVKVIOQESYTY 289

QY 240 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIETFCRIH 299
DB 290 KDPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIETFCRIH 349

QY 300 QCISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 359
DB 350 QCISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 409

QY 360 KSLSRQSLMANNIEKLNQNSRSEAPNWTODSGFY 396
DB 410 KSLSRQSLMANNIEKLNQNSRSEAPNWTODSGFY 446

RESULT 12
Q9CT23 PRELIMINARY; PRT; 422 AA.
AC Q9CT23;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2610017H09 product:mammary tumor integration
DE site 6, full insert sequence. (fragment).
GN Name=Eif386;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK011436; BAB27621.1; -.
DR MGI; MGI:99257; Eif386.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
FT NON TER 422
FT SEQUENCE 422 AA; 49568 MW; A05157DBC6D375B CRC64;
SQ
Query Match 94.0%; Score 1946; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 9.6e-136;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDFAMDVYKNLYSD-DIPHALREKRTTVVAOLKQLOAETEPVVKMFEDPETTRQMOSTR 60
DB 50 MYDFAMDVYKNLYSD-DIPHALREKRTTVVAOLKQLOAETEPVVKMFEDPETTRQMOSTR 109

QY 61 GRMLFDYLADKHGRFQRYELDTLYRYAKFQYECGNYSGAAEYLYFRVLVPAVDNRNALSS 120
DB 110 GRMLFDYLADKHGRFQRYELDTLYRYAKFQYECGNYSGAAEYLYFRVLVPAVDNRNALSS 169

QY 121 WGLKLAISEILMQNWDAAEMDLTRLKETIDNNVSSPLQSLQORTWLIHWSLFFVFNHPKG 180
DB 170 WGLKLAISEILMQNWDAAEMDLTRLKETIDNNVSSPLQSLQORTWLIHWSLFFVFNHPKG 229

QY 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKQVRKROVLKDLVKVIOQESYTY 240
DB 230 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKQVRKROVLKDLVKVIOQESYTY 289

QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 349

QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 360
DB 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIKT 409
```

```
Db 350 CISINMLADKLNMTPEAEERWVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIEKTK 409
QY 361 SLSFRSOMLANNI 373
Db 410 SLSFRSOMLANNI 422
RESULT 13
Q6TH16
ID 06TH16 PRELIMINARY; PRT; 448 AA.
AC Q6TH16;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 3, subunit 6 48kDa.
GN Names-E1F3S6; ORFNames-zgc:63821;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Kidney marrow;
RA Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,
RA Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zou L.I.,
RA Kanki J.P., Look A.T., Chen Z.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY398340; AAQ97773.1; -
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR00717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
KW Initiation factor.
SQ
SEQUENCE 448 AA; 52672 MW; 190C9FD4CEB77803 CRC64;
Query Match 86.1%; Score 1783.5; DB 2; Length 448;
Best Local Similarity 84.6%; Pred. No. 1.1e-123;
Matches 346; Conservative 24; Mismatches 16; Indels 23; Gaps 3;
QY 1 MYDFAMDVKNLYSD-DIPALREKETTVAQLKQLQAEATERIVKMFEDPETTRQMSTR 59
Db 50 MYDFAMDVYRNLPDPKEIPNSUREKETTVAQLKQLQSETEPIVKVFEDPETTRQMSTR 109
QY 60 DGRMLFDYLADKHGFRQYEDLTLYRYAKFYQECGNYSGAAEYLYFPRVLVPAIDRVALGS 119
Db 110 DGRMLFDYLADKHGFRQYEDLTLYRYAKFYQY-----VWELLIVQRSTCTSSRS 159
QY 120 -----LWGLKASEILMQNWDAAEMEDLTRKETIDNNVSSPLQSLQOQRTWLH 167
Db 160 GPCHRQKRAELLWGKLAISEILMQNWEAAMEDLTRLETDNNVSSPLQSLQOQRTWLH 219
QY 168 WSLFVFFNHPKGRDNIIDFLQPOYLNAIQWCPHILRYLTAVITNKKVRKRVOLKD 227
Db 220 WSLFVFFNHPKGRDNIIEFLQPOYLNAIQWCPHILRYLTAVITNKKVRKRVOLKD 279
QY 228 LVKVLQOESYTYKDPITEFEVCLYVNFDFDGAQKKLRCECVLNDFFLVACLEDPIENA 287
Db 280 LVKVLQOESYTYKDPITEFEVCLYVNFDFDGAQKKLRCECVLNDFFLVACLEDPIENA 339
QY 288 RLFIETFCRIHQCSINMLADKLNMTPEAEERWVNLIRNARLDKIDSKLGHVVMGN 347
Db 340 RLFIETFCRIHQCSINMLADKLNMTPEAEERWVNLIRNARLDKIDSKLGHVVMGN 399
QY 348 AVSPYQOVIEKTKSLSFRSOMLANNIEKKLNNSRSEAPNWTQDSGFY 396
Db 400 AISPYQOVIEKTKSLSFRSOMLANNIEKQSNANRNTPNWAQAGFY 448
```

RESULT 14
IF36_DROME

```
ID IF36_DROME STANDARD; PRT; 435 AA.
AC Q774I0; Q9VVA2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Probable eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48) (eIF3e).
GN Name-Int6; ORFNames-CG9677;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Miyazaki S., Diella F., Callahan D., Sullivan D., Callahan R.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=Berkely; TISSUE=Ovary;
RX MEDLINE=20196012; PubMed=10731138; DOI=10.1126/science.287.5461.2222;
RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,
RA Stapleton M., Harvey D.A.;
RT "A Drosophila complementary DNA resource.";
RL Science 287:2222-2224 (2000).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner T.K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [4]
GENOME REANNOTATION.
RP MEDLINE=22426069; PubMed=12537572;
RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
```

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bertencourt B.R., Celiniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.,
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.",
RL Genome Biol. 3:RESEARCH0083.22(2002).
CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of
CC methionyl-tRNA and mRNA (by similarity).
CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits (by
CC similarity).
CC -!- SIMILARITY: Belongs to the eIF36 family.
CC -!- SIMILARITY: Contains 1 PCI domain.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U9162; AAC62307.1; -.
CC EMBL; AF132551; AAD27850.1; -.
CC EMBL; AE003526; AAF49412.1; -.
CC Flybase; FBgn0025582; Int6.
CC InterPro; IPR000717; PCI.
CC Pfam; PF01399; PCI; 1.
CC SMART; SM00088; PINT; 1.
CC Initiation factor; Protein biosynthesis.
CC SEQUENCE 435 AA; 51162 MW; 6F79D32FBA3EA711 CRC64;

Query Match 58.9%; Score 1220.5; DB 1; Length 435;
Best Local Similarity 59.7%; Pred. No. 4.9e-82;
Matches 233; Conservative 65; Mismatches 83; Indels 9; Gaps 5;
QY 1 MVDFAVDYKYL-YSDDIPHALREKRTTVVAOLKQLAETEPVIVKMFDPETTRQVSTR 59
DB 50 MIDYTMTRKRLNLSQEMPEELVQRAEYVATLKLQNEVAPIMKA---TDILKNGESMK 106
QY 60 DGRMLFDYLDKHFQRYEYLDLTLYRYAKFOYECGNYSGAAEYLYFPRVLVPATDRNALSS 119
DB 107 DSKTFNALQDYNFKEVHESAYKLAKLYECGNYQESTSYLYFCLIVMSPNCKNLYNV 166
QY 120 LGKGLASEILMKNQWDAEMDLTRLKETIDNNSSVSPLOQRTWLIHWSLVFFNHPKG 179
DB 167 LGKGLAAEILTNWNTALEDLTRLRDYIDNANFST-IQALQRTWLIHWSLVFFNHPKG 225
QY 180 RDNIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIQOESYTY 239
DB 226 RLIIEMFLYKPLYLNAIQTMCPHIMRYLATAVIN---RTRNALKDLIKVIOQESYTY 282
QY 240 KDPITEFVCLVYNFDFGAKKLRCESVLVNDFLVACLEDFIENARLFIFETPCRIH 299
DB 283 RDPITEFVCLVYNFDFGAKKLRCESVLVNDFLVACLEDFIENARLFIFETPCRIH 342
QY 300 QCISINMLADKLMNTPPEAERWVNLIRNARLDKIDS KLGHVVMGNNAVSPYQVIEKT 359
DB 343 QCITISMLADKLMNTPPEAERWVNLIRNARLDKIDS KLGHVVMGNNAVSPYQVIEKT 402
QY 360 KSLSFQSMAMNIEKLNQNSRSE-APNW 398
DB 403 DLSMRSEHLAGLIERKSKQKQNGESADSW 432

RESULT 15

Q7QIL8 PRELIMINARY; PRT; 462 AA.
AC Q7QIL8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE AGCP326 (Fragment).
GN Name=agCG57001; ORFNames=ENSANGG00000019251;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAB01008807; EAA04727.1; -.
CC InterPro; IPR000717; PCI.
CC Pfam; PF01399; PCI; 1.
CC NON_TER 1
SQ SEQUENCE 462 AA; 53647 MW; 835173A6776AA4E0 CRC64;
Query Match 56.3%; Score 1165.5; DB 2; Length 462;
Best Local Similarity 58.0%; Pred. No. 6.3e-78;
Matches 232; Conservative 56; Mismatches 83; Indels 29; Gaps 7;
QY 1 MVDFAVDYKYL-YSDDIPHALREKRTTVVAOLKQLAETEPVIVKMFDPETTRQVSTR 59
DB 79 MVDYKHDIRERLAMDVKVHDELAQGRANVLATLKELOAEVAPLMKMBE---LKNPDSK 135
QY 60 DGRMLFDYLDKHFQRYEYLDLTLYRYAKFOYECGNYSGAAEYLYFPRVLVPATDR 114
DB 136 DSKSVIHAL-----QOTLDYDIILSAQKLAKLYECGNYNDSLYLVCMVMEPNDK 188
QY 115 NALSLGKGLASEILMKNQWDAEMDLTRLKETIDNNSSVSPLOQRTWLIHWSLVFF 174
DB 189 NYLGLVWGLKLAIVEILTLNWTALDELTRLRDFIENYF-SPIQVLRQARWLHWSLVFF 247
QY 175 NHPKGRDNIIDLFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIQO 234
DB 248 NHGKGRDLIIDMFLYKPOYLNAIQTMCPHILRYLATAVIN---RGRNALKDLIKVIOQ 304
QY 235 ESYTYKDPITEFVCLVYNFDFGAKKLRCESVLVNDFLVACLEDFIENARLFIFET 294
DB 305 ESYTYRDPITEFVCLVYNFDFGAKKLRCESVLVNDFLVACLEDFIENARLFIFET 364
QY 295 FCRIHQCISINMLADKLMNTPPEAERWVNLIRNARLDKIDS KLGHVVMGNNAVSPYQO 354
DB 365 FCRIHQCITIGMLADKLMNTPPEAERWVNLIRNARLDKIDS KLGHVVMGNNAVSPYQO 424
QY 355 VIEKTSKLSFRSOMLAMNIEKLNQNSRSEAPNWTQDSG 394
DB 425 LIVEKIDSLSVRSEALTLIAVERKHAK-----TQEAG 455

Search completed: March 8, 2005, 20:07:17
Job time : 175 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:07:23 ; Search time 136 Seconds
(without alignments)
958.101 Million cell updates/sec

Title: US-10-783-415-4
Perfect score: 2071
Sequence: 1 MVDFAVDVKNLYSDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 32904822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pcp:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pcp:*
18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pcp:*
19: /cgn2_6/ptodata/2/pubpaa/US10G_PUBCOMB.pcp:*
20: /cgn2_6/ptodata/2/pubpaa/US10H_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2071	100.0	396	10	US-09-858-152A-2
2	2071	100.0	396	10	US-09-858-152A-4
3	2071	100.0	396	16	US-10-408-765A-2064
4	2071	100.0	396	16	US-10-783-415-2
5	2071	100.0	396	16	US-10-783-415-4
6	2071	100.0	445	13	US-10-087-192-558
7	1279.5	61.8	320	13	US-10-087-192-558
8	1087.5	52.5	439	16	US-10-767-701-45301
9	1085.5	52.4	439	16	US-10-437-963-188307
10	1064	51.4	437	15	US-10-424-599-266298
11	952.5	46.0	444	16	US-10-437-963-188308
12	564	27.2	117	15	US-10-264-049-4238
13	555	26.8	187	15	US-10-425-114-61814

14	391.5	18.9	130	15	US-10-424-599-217317	Sequence 217317,
15	260	12.6	166	16	US-10-437-963-188446	Sequence 188446,
16	223	10.8	45	9	US-09-864-761-42338	Sequence 42338, A
17	211	10.2	42	9	US-09-864-761-40729	Sequence 40729, A
18	211	10.2	42	9	US-09-864-761-42138	Sequence 42138, A
19	173	8.4	112	14	US-10-106-698-6920	Sequence 6920, Ap
20	135	6.5	61	15	US-10-424-599-213418	Sequence 213418,
21	114	5.5	381	15	US-10-424-599-225049	Sequence 225049,
22	112	5.4	567	13	US-10-087-192-1170	Sequence 1170, Ap
23	110	5.3	858	15	US-10-425-114-72451	Sequence 72451, A
24	108	5.2	147	15	US-10-425-114-49583	Sequence 49583, A
25	107.5	5.2	641	16	US-10-444-693C-36	Sequence 36, Appl
26	106.5	5.1	1186	15	US-10-282-122A-60346	Sequence 60346, A
27	106	5.1	1146	15	US-10-282-122A-66407	Sequence 66407, A
28	106	5.1	1575	16	US-10-408-765A-704	Sequence 704, App
29	105	5.1	1323	15	US-10-389-566-1642	Sequence 1642, Ap
30	104.5	5.0	681	15	US-10-282-122A-71657	Sequence 71657, A
31	104.5	5.0	684	15	US-10-104-047-2418	Sequence 2418, Ap
32	103.5	5.0	618	14	US-10-203-860-24	Sequence 24, Appl
33	103.5	5.0	742	14	US-10-203-860-2	Sequence 2, Appli
34	103.5	5.0	742	14	US-10-308-448-11	Sequence 11, Appl
35	103.5	5.0	742	15	US-10-341-434-85	Sequence 85, Appl
36	102.5	4.9	806	16	US-10-437-963-149528	Sequence 149528,
37	101.5	4.9	872	14	US-10-326-040-7	Sequence 7, Appli
38	101	4.9	1039	15	US-10-369-493-16475	Sequence 16475, A
39	100.5	4.9	87	15	US-10-424-599-168200	Sequence 168200,
40	100.5	4.9	872	15	US-10-424-599-219651	Sequence 219651,
41	100.5	4.9	1020	15	US-10-282-122A-58016	Sequence 58016, A
42	100	4.8	1575	17	US-10-741-600-1548	Sequence 1548, Ap
43	100	4.8	1575	17	US-10-741-600-1549	Sequence 1549, Ap
44	100	4.8	1640	17	US-10-741-600-1550	Sequence 1550, Ap
45	99.5	4.8	4563	16	US-10-741-601-432	Sequence 432, App

ALIGNMENTS

RESULT 1
US-09-858-152A-2
; Sequence 2, Application US/09858152A
; Publication No. US2003004419A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE
; APPLICANT: SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858,152A
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO. 2
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Murine INT6
US-09-858-152A-2

Query Match 100.0%; Score 2071; DB 10; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-179; Indels 0; Gaps 0;
Matches 396; Conservative 0; Mismatches 0;
QY 1 MVDFAVDVKNLYSDIPHALEKRTTVVAQKQLQAEETPIVKMFEDPETTRQMSTRD 60
DB 1 MVDFAVDVKNLYSDIPHALEKRTTVVAQKQLQAEETPIVKMFEDPETTRQMSTRD 60
QY 61 GRMLFDYLDKGFQREYLDLTYRYAKFQYECGNYSAAEYLYFFRVLPATDNRNALSSL 120
DB 61 GRMLFDYLDKGFQREYLDLTYRYAKFQYECGNYSAAEYLYFFRVLPATDNRNALSSL 120
QY 121 WGLKASEILMQWDAAMEDLTRKETIDNNSVSSQLSQORTWLHWSLVFFNPKGR 180

```
Db 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 240
Db 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 2
US-09-858-152A-4
; Sequence 4, Application US/09858152A
; Publication No. US2003004419A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE
; APPLICANT: SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitia, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858,152A
; CURRENT FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-152A-4

Query Match 100.0%; Score 2071; DB 10; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-179;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAQLKQLQAETEPVVKMPDPETTRQMOSTRD 60
Db 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAQLKQLQAETEPVVKMPDPETTRQMOSTRD 60
QY 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLFFRVLVPATDRNALSSL 120
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
Db 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 240
Db 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
```

```
Db 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 3
US-10-408-765A-2064
; Sequence 2064, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Faby, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2064
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2064

Query Match 100.0%; Score 2071; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-179;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAQLKQLQAETEPVVKMPDPETTRQMOSTRD 60
Db 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAQLKQLQAETEPVVKMPDPETTRQMOSTRD 60
QY 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLFFRVLVPATDRNALSSL 120
Db 61 GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSGAAEYLFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
Db 121 WGLKASEILMQNWDAAEDLTRELKETIDNNSVSSPLQSLQRTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 240
Db 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVQQESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKKLRECVSLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 4
US-10-783-415-2
; Sequence 2, Application US/10783415
; Publication No. US20040141918A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitia, Flamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
```


Query Match 100.0%; Score 2071; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-179; Indels 0; Gaps 0;
Matches 396; Conservative 0; Mismatches 0;
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
FILE REFERENCE: 4239-59122
CURRENT APPLICATION NUMBER: US/10/783,415
CURRENT FILING DATE: 2004-02-19
PRIOR APPLICATION NUMBER: 09/858,152
PRIOR FILING DATE: 2001-05-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 396
TYPE: PRT
ORGANISM: Murine INT6
US-10-783-415-4

Query Match 100.0%; Score 2071; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-179; Indels 0; Gaps 0;
Matches 396; Conservative 0; Mismatches 0;
QY 1 MVDPMVYKNLYSDDI PHALREKRTTVAQKQLOAETEPVVKMFEDPETTRQMOSTRD 60
DB 1 MVDPMVYKNLYSDDI PHALREKRTTVAQKQLOAETEPVVKMFEDPETTRQMOSTRD 60
QY 61 GRMLFDYADKHGFRQYLDLTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYADKHGFRQYLDLTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240
DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLANTPPEAERWIINLRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 301 CISINMLADKLANTPPEAERWIINLRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
DB 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 5

US-10-783-415-4
Sequence 4, Application US/10783415
Publication No. US20040141918A1
GENERAL INFORMATION:
APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
HUMAN SERVICES
APPLICANT: Marchetti, Antonio
APPLICANT: Buttitta, Fiamma
APPLICANT: Smith, Gilbert H.
APPLICANT: Callahan, Robert
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
FILE REFERENCE: 4239-59122
CURRENT APPLICATION NUMBER: US/10/783,415
CURRENT FILING DATE: 2004-02-19
PRIOR APPLICATION NUMBER: 09/858,152
PRIOR FILING DATE: 2001-05-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4
LENGTH: 396
TYPE: PRT
ORGANISM: Homo sapiens
US-10-783-415-4

Query Match 100.0%; Score 2071; DB 16; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-179; Indels 0; Gaps 0;
Matches 396; Conservative 0; Mismatches 0;
QY 1 MVDPMVYKNLYSDDI PHALREKRTTVAQKQLOAETEPVVKMFEDPETTRQMOSTRD 60
DB 1 MVDPMVYKNLYSDDI PHALREKRTTVAQKQLOAETEPVVKMFEDPETTRQMOSTRD 60
QY 61 GRMLFDYADKHGFRQYLDLTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYADKHGFRQYLDLTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240
DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLANTPPEAERWIINLRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 301 CISINMLADKLANTPPEAERWIINLRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
QY 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
DB 361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 6

US-10-087-192-558
Sequence 558, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: CANCER
FILE REFERENCE: 529452000122
CURRENT APPLICATION NUMBER: US/10/087,192
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 558
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens
US-10-087-192-558

Query Match 100.0%; Score 2071; DB 13; Length 445;
Best Local Similarity 100.0%; Pred. No. 8.5e-179; Indels 0; Gaps 0;
Matches 396; Conservative 0; Mismatches 0;
QY 1 MVDPMVYKNLYSDDI PHALREKRTTVAQKQLOAETEPVVKMFEDPETTRQMOSTRD 60
DB 50 MVDPMVYKNLYSDDI PHALREKRTTVAQKQLOAETEPVVKMFEDPETTRQMOSTRD 109
QY 61 GRMLFDYADKHGFRQYLDLTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
DB 110 GRMLFDYADKHGFRQYLDLTLYRYAKFOYECGNYSGAAEYLYFFRVLVPATDRNALSSL 169
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQOORTWLHWSLFFVFNHPKGR 229

QY 181 DNIIDFLYQPOYLNAIQMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTYK 240
Db 230 DNIIDFLYQPOYLNAIQMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTYK 289
QY 241 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300
Db 290 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 349
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNVSPYQOVIETK 360
Db 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNVSPYQOVIETK 409
QY 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 410 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 7

US-10-087-192-555
; Sequence 555, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 555
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-555

Query Match 61.8%; Score 1279.5; DB 13; Length 320;

Best Local Similarity 66.4%; Pred. No. 2.6e-107;

Matches 263; Conservative 0; Mismatches 0; Indels 133; Gaps 1;

QY 1 MYDFAMDVYKNLYSDDIIPHALREKRTTVVAQLKOLQAETEPVVKMFPEDPETTRQWQSTRD 60
Db 58 MYDFAMDVYKNLYSDDIIPHALREKRTTVVAQLKOLQAETEPVVKMFPEDPETTRQWQSTR- 116
QY 61 GRMLFDYLDKRGFRQYDLTLRYAKFOYECGNYSAGAEYLFFPRVLVPATDRNALSSL 120
Db 117 ----- 116
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSPLOSQRTWLIIHWSLFFVFNHPKGR 180
Db 117 ----- 116
QY 181 DNIIDFLYQPOYLNAIQMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTYK 240
Db 117 -----YLNAIQTMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTYK 164
QY 241 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300
Db 165 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 224
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNVSPYQOVIETK 360
Db 225 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNVSPYQOVIETK 284
QY 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 285 SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 320

RESULT 8

US-10-767-701-45301
; Sequence 45301, Application US/10767701
; Publication No. US20040172884A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45301
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C1020_1.pep
US-10-767-701-45301

Query Match 52.5%; Score 1087.5; DB 16; Length 439;

Best Local Similarity 53.2%; Pred. No. 1e-89;

Matches 205; Conservative 84; Mismatches 77; Indels 19; Gaps 6;

QY 1 MYDFAMDVYKNLYSDDIIPHALREKRTTVVAQLKOLQAETEPVVKMFPEDPETTRQWQSTR 59
Db 50 MYDYAMDIHKSILHGTHDDVPEDMVKRAEVSRLSRLEEAAPLVAFLQNPQLVQELRP-- 107
QY 60 DGRMLFDYLDKRGFRQYDLTLRYAKFOYECGNYSAGAEYLFFPRVLVPATDRNALSS 119
Db 108 DKQYNIHMLQERYQIGPQIEALYQAKFOYECGNYSAGADLYQYRALTCTNSR-SVSA 166
QY 120 LMGKLAASEILMQNWDAAEMDLTRLKETIDNNSVSPLOSQRTWLIIHWSLFFVFNHPK 179
Db 167 LMGKLAASEILMQNWDVALEELNRLKEIIDSKNFSPLQNLQNIWLMHWALSIFFNHENG 226
QY 180 RDNIIIDFLYQPOYLNAIQMCPHILRYLTAVITNKDKRRQVLKDLVKVIOQESYTY 239
Db 227 RNGIIDLFP-QDRLYNALQTNNAHLRLYLATAVVVYK---RRNNMLKELIKVIOQEQHY 282
QY 240 KDPITFEVCLYVNFDFDGAQKLRCESVLVNDFFL-----VACLEDFTENAR 288
Db 283 KDPITFEVCLYVNFDFDGAQKLRCESVLVNDFFLGKRIBEGNFVTVPLRDEFLENAR 342
QY 289 LFIFETFCRIHQICISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNN 348
Db 343 LFIFETFCRIHRCIDISMLSQKLNMSYDEAEELWIMNLVRSKLDARIDSVSGTLIMTTNH 402
QY 349 VSPYQOVIETKSLSPRSQMLAMNI 373
Db 403 VNVHEQIIENLKGKLNMTYMLAKNI 427

RESULT 9

US-10-437-963-188307
; Sequence 188307, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

```
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188307
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84924C.1.pep
US-10-437-963-188307

Query Match          52.4%; Score 1085.5; DB 16; Length 439;
Best Local Similarity 54.3%; Pred. No. 1.6e-89;
Matches 209; Conservative 78; Mismatches 79; Indels 19; Gaps 6;

QY      1 MVDFAVDVYKNLY-SDDIPHALREKTTTVAQKQLQAEPEPIVKMFEDPETTRQMOSTR 59
DB      50 MVDYAMDIHKSLSHDTDEVDDMVRRRTDVVSRULDEATAPIVSLQNLQQLVRLP-- 107

QY      60 DGRMLFDYLADKHGFRQEVLDLTYRYAKFOYECGNYSGAAEYLYFFRVLPVATDRNALSS 119
DB      108 DKQYNLMLQDRFOIGPEIETUYAKQFDCGNTSDAAVLYQYRALCTNSER-SLSA 166

QY      120 LMGKLAASEILMONWDALDELNRLKEIIDSXKNSFSPNLQNLQNRILMLMWSIFIFFNHENG 179
DB      167 LMGKLAASEILMONWDALDELNRLKEIIDSXKNSFSPNLQNLQNRILMLMWSIFIFFNHENG 226

QY      180 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLVKVIQESYTY 239
DB      227 RNGIIDLFP-FQDKYLNAIQTNAPHLRLYLATAVVVVK--RRRNMLKELIKVIOQBSHY 282

QY      240 KDPITEFVECLYVNFDFGAQKQKRECESVLNDPFL-----VACLEDFTENAR 288
DB      283 KDPITEFVECLYVNFDFGAQKQKRECEVILNDPFLGKRVEESNFTVPLRDEFTENAR 342

QY      289 LFIFETFCRIHQICISINMLADKLNTPEEAERWIVNLIRNARDKIDSKLGHVVMGNNA 348
DB      343 LFIFETFCRIHRCIDIGMLSQKLMNSYDEAELWIMLVNSKLDKIDSVGTLINTNH 402

QY      349 VSPYQVQVIEKTKSLSPRSQMLANWI 373
DB      403 VNIHQEVIESLKNLNRMTFLAKNI 427

RESULT 10
US-10-424-599-266298
; Sequence 266298, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 266298
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_82489C.1.pep
US-10-424-599-266298

Query Match          51.4%; Score 1064; DB 15; Length 437;
Best Local Similarity 52.4%; Pred. No. 1.4e-87;
Matches 207; Conservative 82; Mismatches 86; Indels 20; Gaps 7;

QY      1 MVDFAVDVYKNLY-SDDIPHALREKTTTVAQKQLQAEPEPIVKMFEDPETTRQMOSTR 59
DB      50 MVDYAMDIHKSLSYHTDEVDDMDVRAEVVARLKSLEEAAPLVAFQNEAAVQELRA--- 107

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188307
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84924C.1.pep
US-10-437-963-188307

Query Match          46.0%; Score 952.5; DB 16; Length 444;
Best Local Similarity 49.9%; Pred. No. 1.8e-77;
Matches 197; Conservative 68; Mismatches 93; Indels 37; Gaps 8;

QY      1 MVDFAVDVYKNLYSD-----DIPHALREKTTTVAQKQLQAEPEPIVKMFED----P 49
DB      53 MVDYAMDTKHLHGDTDDVVVPVDDMVRRHEVTRLGALAAPAPIVSALKNHHLGP 112

QY      50 ETTRQMOSTRGRMLFDYLADKHGFRQEVLDLTYRYAKFOYECGNYSGAAEYLYFFRVLV 109
DB      113 DKEHNI-----RMLH-----ERFOIGPQDQIEALYQYAKFQFCGNYPDAENLHYRALC 162

QY      110 PATDRNALSSLWGLASRILMONWDAAMEDLTLKXETIDNNSSVSPSLQSLQORTWLIHWS 169
DB      163 TYSER-SUSAQMGKLSAEILNNWDAEELNCLKEMIDSKNSSSSPNLQNLRIWLMHCS 221

QY      170 LFVFFNHKPGRDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLV 229
DB      222 IFIFNHGNGSGIIDLFP-FQDKYLNAIQTDAPHLRLYLAAAVVN--RRRNMLKELV 277
```

```
QY      60 DGRMLFDYLADKHGFRQEVLDLTYRYAKFOYECGNYSGAAEYLYFFRVLPVATDRNALSS 119
DB      108 DKQYNLQMLNDRYQIGPAQIEALYQYAKFQFCGNYSGAADYLYQYRALCTNSER-SLSA 166

QY      120 LMGKLAASEILMONWDAAMEDLTLKXETIDNNSSVSPSLQSLQORTWLIHWSLFFVFNHHPKG 179
DB      167 LMGKLAASEILMONWDIALBELNRLKEIIDSXKNSFSPNLQNLQNRILMLMWSLFFVFNHNDG 226

QY      180 RDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLVKVIQESYTY 239
DB      227 RTQIIDLF-NQDKYLNAIQTSAPHLRLYLATAPIVVK--RRRPQKDFIKVIOQBSHY 282

QY      240 KDPITEFVECLYVNFDFGAQKQKRECESVLNDPFL-----VACLEDFTENAR 288
DB      283 KDPITEFLACVYVNFDFGAQKQKRECEVILNDPFLGKRVEESNFTVPLRDEFTENAR 342

QY      289 LFIFETFCRIHQICISINMLADKLNTPEEAERWIVNLIRNARDKIDSKLGHVVMGNNA 348
DB      343 LFIFETFCRIHQRIDGVLAELKLNLYEEAERWIVNLIRSKLDKIDSETGAVIMEPNH 402

QY      349 VSPYQVQVIEKTKSLSPRS-QMLAMNIEKKLNQNSR 382
DB      403 LNVYEQLIDHTKALNGRTVTKLVSQLLEHAQAQOTTR 437

RESULT 11
US-10-437-963-188308
; Sequence 188308, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188308
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84925C.1.pep
US-10-437-963-188308

Query Match          46.0%; Score 952.5; DB 16; Length 444;
Best Local Similarity 49.9%; Pred. No. 1.8e-77;
Matches 197; Conservative 68; Mismatches 93; Indels 37; Gaps 8;

QY      1 MVDFAVDVYKNLYSD-----DIPHALREKTTTVAQKQLQAEPEPIVKMFED----P 49
DB      53 MVDYAMDTKHLHGDTDDVVVPVDDMVRRHEVTRLGALAAPAPIVSALKNHHLGP 112

QY      50 ETTRQMOSTRGRMLFDYLADKHGFRQEVLDLTYRYAKFOYECGNYSGAAEYLYFFRVLV 109
DB      113 DKEHNI-----RMLH-----ERFOIGPQDQIEALYQYAKFQFCGNYPDAENLHYRALC 162

QY      110 PATDRNALSSLWGLASRILMONWDAAMEDLTLKXETIDNNSSVSPSLQSLQORTWLIHWS 169
DB      163 TYSER-SUSAQMGKLSAEILNNWDAEELNCLKEMIDSKNSSSSPNLQNLRIWLMHCS 221

QY      170 LFVFFNHKPGRDNIIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVKRRQVLDKLV 229
DB      222 IFIFNHGNGSGIIDLFP-FQDKYLNAIQTDAPHLRLYLAAAVVN--RRRNMLKELV 277
```

```
QY 230 KVIQESYTKDPITEFVECLVNFDFGAQKKECESVLVNDPFL-----VA 278
Db 278 KVIQEQSHYKDPITEFVECLVNFDFGAQKKECESVLVNDPFLGKRKEGNSITVP 337
QY 279 CLEDFIENARLFIPTFCRIHQCSINMLADKLNMTPEAEARWVNLIRNARLDKIDSK 338
Db 338 LRDEPLENARLFIPTFCRIHRSIDIGMLSQKLNRYDEGELWMLNVLVNSKLDKIDSV 397
QY 339 LGHVVGNGNVSFYOQVTEKTSLSFRSQMLAMNI 373
Db 398 SGTLMTTHNVNHEQFIESLKNLNMRTSMLAKNI 432

RESULT 12
US-10-264-049-4238
; Sequence 4238, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 4238
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-4238

Query Match 27.2%; Score 564; DB 15; Length 117;
Best Local Similarity 95.6%; Pred. No. 41e-43;
Matches 108; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 122 GKXSEILMOWDAAWEDLTRKETIDNNSVSSPLQSLQORTWLHWSLVFVFNHPKGRD 181
Db 1 6KXSEILMOWDAAWEDLTRKETIDNNSVSSPLQSLQORTWLHWSLVFVFNHPKGRD 60

QY 182 NIIDLFLYQPYLNAIQTMCPIHLRYLTAVITNKDVRKRRQVLKDLVKVIOQ 234
Db 61 NIIDLFLYQPYLNAIQTMCPIHLRYLTAVITNKDVRKRRQVLKDLGKVFNR 113

RESULT 13
US-10-425-114-61814
; Sequence 61814, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61814
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Zea mays
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73157G05_FLI.pep
US-10-425-114-61814

Query Match 26.8%; Score 555; DB 15; Length 187;
Best Local Similarity 57.3%; Pred. No. 5.2e-42;
Matches 102; Conservative 33; Mismatches 25; Indels 18; Gaps 3;

QY 77 EYLDTRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSLWGLKASEILMOWDAA 136
Db 23 DQIEALYQYAFQFECGNYSGADYLYOYGALCTNSER-VSALWGLKAAEILMOWDIA 81

QY 137 MEDTLRLKETIDNNSVSSPLQSLQORTWLHWSLVFVFNHPKGRDNIIDLFLYQPYLNA 196
Db 82 LEELNRLKEIIDSKNFSPLNQLNRIWLHWFALFIPFNHIR-----YLNA 127

QY 197 IQTWCPHILRYLTAVITNKDVRKRRQVLKDLVKVIOQESYTKDPITEFVECLVNF 254
Db 128 IQTNAHLIRYLATAVVVVK---RRNMLKELIKVIOQEHSHYKDPVTFEULECLVNY 182

RESULT 14
US-10-424-599-217317
; Sequence 217317, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 217317
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(130)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MET3847_38264C.1.pep
US-10-424-599-217317

Query Match 18.9%; Score 391.5; DB 15; Length 130;
Best Local Similarity 64.2%; Pred. No. 2e-27;
Matches 70; Conservative 19; Mismatches 19; Indels 1; Gaps 1;

QY 79 LDTLYRYAKFQYECGNYSGAAEYLYFFRVLPATDRNALSSLWGLKASEILMOWDAAE 138
Db 12 IEALYQYAKFQYECGNYSGAADYLYOYRALCTNSER-SLSALWGLKAAEILMOWDIALE 70

QY 139 DLTRLKETIDNNSVSSPLQSLQORTWLHWSLVFVFNHPKGRDNIIDLFL 187
Db 71 ELNRLKEIIDSKNFSPLNQVQSRIWLHWSLVFVFNHNDNGTXIIDLF 119

RESULT 15
US-10-437-963-188446
; Sequence 188446, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
```

Search completed: March 8, 2005, 20:19:52
Job time : 138 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 19:49:22 ; Search time 41 Seconds
(without alignments)
929.313 Million cell updates/sec

Title: US-10-783-415-4
Perfect score: 2071
Sequence: 1 MVDFAVDYKKNLYSDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2071	100.0	396	2	I75615 mammary tumor inte
2	994.5	48.0	432	2	T33118 hypothetical prote
3	903	43.6	418	2	T45807 translation initia
4	775	37.4	501	2	T40585 translation initit
5	760	36.7	501	2	T50488 translation initit
6	116	5.6	458	2	S56816 GTPase-activating
7	114	5.5	1158	2	T43519 condensin complex
8	114	5.5	2121	2	T27406 hypothetical prote
9	109.5	5.3	446	2	H97754 lipid-A-disacchari
10	108	5.2	481	2	G97089 gamma-glutamylcyst
11	106.5	5.1	1186	2	AD1300 Smc protein essent
12	106	5.1	1146	2	C83304 hypothetical prote
13	105.5	5.1	1186	2	AD1672 Smc protein essent
14	105	5.1	668	2	S64123 hypothetical prote
15	103.5	5.0	742	2	JC7595 scavenger receptor
16	103	5.0	1173	2	T30308 rxaA protein - Lac
17	102	4.9	2346	2	T13829 Tpr homolog - frui
18	101.5	4.9	622	2	D84493 probable retroelem
19	100	4.8	569	1	S62851 hypothetical prote
20	99.5	4.8	300	2	T20627 hypothetical prote
21	99.5	4.8	341	2	T20626 hypothetical prote
22	99.5	4.8	732	2	T32757 hypothetical prote
23	99	4.8	519	2	T27880 hypothetical prote
24	99	4.8	1060	2	F88710 protein ColG5.4 [i
25	99	4.8	1070	2	G84982 exodeoxyribonuclea
26	99	4.8	1079	2	T30996 hypothetical prote
27	99	4.8	1927	2	T25604 hypothetical prote
28	99	4.8	2541	2	T29340 hypothetical prote
29	98.5	4.8	1367	2	H82874 conserved hypothet

30	98.5	4.8	2663	1	S28261 centromere protein
31	98.5	4.8	4563	1	LPH0B apolipoprotein B-1
32	98	4.7	760	2	T28224 ORF MSV063 probabl
33	98	4.7	979	2	JQ0894 P115 protein - Myc
34	97.5	4.7	509	2	B90573 hypothetical prote
35	97.5	4.7	761	2	T03719 probable thyroid r
36	97.5	4.7	1156	2	E69444 chromosome segrega
37	97.5	4.7	1792	2	T20363 hypothetical prote
38	97	4.7	401	2	AI2255 two-component sens
39	97	4.7	716	2	T02311 hypothetical prote
40	97	4.7	1301	2	S51323 SAC3 protein - yea
41	96.5	4.7	787	2	H90543 conserved hypothet
42	96.5	4.7	897	2	G02529 dynein heavy chain
43	96.5	4.7	4644	1	A38905 dynein heavy chain
44	96	4.6	383	2	S6729 probable membrane
45	96	4.6	759	2	T42652 hypothetical prote

ALIGNMENTS

RESULT 1

I75615
mammary tumor integration site 6 oncogene protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999
C:Accession: I75615; I75614
R:Marchetti, A.; Buttitta, F.; Miyazaki, S.; Gallahan, D.; Smith, G.H.; Callahan, R.
J. Virol. 69, 1932-1938, 1995
A:Title: Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary
A:Reference number: I56914; MUID:95156630; PMID:7853537
A:Accession: I75615
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-396 <MAR1>
A:Cross-references: GB:S75224; NID:g913340; PID:g913341
A:Accession: I75614
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 235-268, 'LKIKFFQSKGIRYVNOAVSGFMEIHRAAVGRPEGHRLGDPLEGNG' <MAR2>
A:Cross-references: GB:S75223; NID:g913338; PIDN:AAC00047.1; PID:g913339
A:Note: hypothetical mutant fusion protein
C:Genetics:
A:Gene: Int-6

Query Match	100.0%;	Score 2071;	DB 2;	Length 396;
Best Local Similarity	100.0%;	Pred. No. 4.9e-147;		
Matches 396;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MVDFAMDVYKKNLYSDDIPHALREKRTTVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD	60	
Db	1	MVDFAMDVYKKNLYSDDIPHALREKRTTVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD	60	
Qy	61	GRMLFDYLADKHGRQEQVLDLYRYAKFQYECGNYSGAAEVLFFRVLPVATDRNALSSL	120	
Db	61	GRMLFDYLADKHGRQEQVLDLYRYAKFQYECGNYSGAAEVLFFRVLPVATDRNALSSL	120	
Qy	121	WGKLASEILMQNDAAMEDLTRLKETIDNNSSVPSLSQLOQRTWLIIHWSLFFVFNHPKGR	180	
Db	121	WGKLASEILMQNDAAMEDLTRLKETIDNNSSVPSLSQLOQRTWLIIHWSLFFVFNHPKGR	180	
Qy	181	DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKQVRKRQVLKDLVKVYIQBSYTYK	240	
Db	181	DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKQVRKRQVLKDLVKVYIQBSYTYK	240	
Qy	241	DPITEFECLVYNFDFGAKKRECESVLVNDFLVACLDFETIENARLFFETFCRIHQ	300	
Db	241	DPITEFECLVYNFDFGAKKRECESVLVNDFLVACLDFETIENARLFFETFCRIHQ	300	
Qy	301	CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDS KLGHVWGNNAVSPVQVIEKTK	360	
Db	301	CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDS KLGHVWGNNAVSPVQVIEKTK	360	

100

[illegible]

A:Molecule type: DNA
A:Residues: 1-1146 <STO>
A:Cross-references: UNIPROT:Q910B0; GB:AE004701; GB:AE004806; PIDN:AG0612
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2732

Query Match 5.1%; Score 106; DB 2; Length 1146;
Best Local Similarity 23.6%; Pred. No. 9.2; Mismatches 54; Gaps 15;
Matches 65; Conservative 50; Indels 106; Indels 54; Gaps 15;

QY 17 IPHALREKRTTVA---QLKQLOAETEPVIMFED--PETTRQMOSTRDGRML--FDYL 68
DB 323 IQHSAGSGKNTIAMLAHQLVELRSVADPLMPQFDSIIIVTDRALDQIARTIKGYD 382

QY 69 ADKHGF---RQBYLDTLYR-----YAKFOY---ECNYSGAAYLYFFRVLVPATDR 114
DB 383 ASIFGHSNDAQELRDYLRGRKIIIVTVQKFPFLEIGDLSGRT-----FALLID--- 433

QY 115 NALSSLMGKLASEILMONWDAAMEDLTRKKTIDNNSVSSPLQS-LQORTWLIHWSLFV 173
DB 434 EAHSSQGGKTTAR--MHEALSGKTDEEAFEDATQAVNTEIEKRIQSRMLSNASYAF 491

QY 174 FHHPKGRDNIIDLFLYQPOLNAIOTMCPHILRYLTAVITNKDVKRRQVLKD--LVKV 231
DB 492 TATPKNR--TSLFGEKTVDDKVEFRSPEELTYTT-----KQAIQEGFILDV 537

QY 232 IQOESYTYKDPITEFVECLYVNFDFGAQ--KKLR 264
DB 538 IAH--YTTLDSFVQAKTVENDPDPFKARALKKIR 570

RESULT 13
AD1672
Smc protein essential for chromosome condensation and partition homolog smc [imported]
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD1672
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simeos, N.; Trierz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative Genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1672
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1186 <GLA>
A:Cross-references: UNIPROT:Q92AK4; GB:AL592022; PIDN:CAC97148.1; PID:G16414419; GSPDB:G
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: smc

Query Match 5.1%; Score 105.5; DB 2; Length 1186;
Best Local Similarity 20.8%; Pred. No. 10;
Matches 90; Conservative 65; Mismatches 177; Indels 107; Gaps 18;

QY 6 MDVYKNLYSDDIPHALREKRTTVAQLKQLOAETEPVIMFEDPETTRQMOSTRDGRMLF 65
DB 452 MEIY-----REVQOTLAKQEAFTGTERALYKH---ETVQMKSRKE---TL 493

QY 66 DYLDKXH-GFRQBYLDTLYRYAKFQYECNYSGAAYLYFFRVLVPATDRNALSSLMGKL 124
DB 494 EELADDYAGFFQGVREVL-----KAKKEIPGIHGVEL-----IDIPAKYQAMETALGAS 545

QY 125 ASILMONWDAAMEDLTRKKTIDNNSVSSPLQSLOORTWLIHWSLFVFNHPKGRDNI 184
DB 546 AQNVVVEDDRVAREATISFLKTKSGRATPLPLSTIQPR-----ELPAATKNAL 593

QY 185 DLFLYQPOY-----LNAIOTMCPHILRYLTAVITNKDVKRRQVLKDVLV----KVIO 233

Db 594 S---NQPAFIALASEVISFDQKQVSPVILNALTGTTILA-KDL-KGANTLARLNFYRIYV 648
QY 234 QES-----YTYKDPITEFVECL-----YVNFDFDGAQ 260
Db 649 LEGDVNAGSGMTGGATKGGKSSILTRKHELQOLAEKLAELNNSSTRELESQVAKDSMS 708
QY 261 KKLRECESVLV-----NDFFLVACLEDFTENARLFFFTETPCRIHQICISINML--ADKL 311
Db 709 KKRELETRVIGENLRLOKEKELKLDRETN-----LERFNKQLQLYDIEKADGSEEL 763
QY 312 NMTPEAEARWVNLNRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETKSLSPRSOMLAM 371
Db 764 NKLLERKETLQEQVEIAKQIEKTEDEETKAMTSSKALES-----KRTADLSSLSLKAQ 818
QY 372 NIEKLNQNSRSEAPNWT 390
Db 819 IAKREQLQSAVEAVDRVT 837

RESULT 14
S64123
Hypothetical protein YGL113w - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein G2980
C:Species: *Saccharomyces cerevisiae*
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 09-Jul-2004
C:Accession: S64123; S64121; S69417
R:Lauguin, G.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64122
A:Accession: S64123
A:Molecule type: DNA
A:Residues: 1-668 <LAU>
A:Cross-references: UNIPROT:P53135; EMBL:Z72635; NID:G1322662; PIDN:CAA96821.1; PID:G13
R:Castagnoli, L.; Paoluzi, S.; Minenkova, O.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64112
A:Accession: S64121
A:Molecule type: DNA
A:Residues: 492-668 <CAS>
A:Cross-references: EMBL:Z72635; MIPS:YGL113w
R:Castagnoli, L.; Paoluzi, S.; Minenkova, O.
submitted to the EMBL Data Library, April 1996
A:Reference number: S69417
A:Accession: S69417
A:Molecule type: DNA
A:Residues: 492-668 <CAW>
A:Cross-references: EMBL:X97644; NID:G1310710; PIDN:CAA66239.1; PID:G1310711
C:Genetics:
A:Gene: SGD:SLD3
A:Cross-references: SGD:S0003081
A:Map position: 7L
C:Superfamily: *Saccharomyces cerevisiae* hypothetical protein YGL113w

Query Match 5.1%; Score 105; DB 2; Length 668;
Best Local Similarity 20.8%; Pred. No. 5.4;
Matches 91; Conservative 72; Mismatches 144; Indels 130; Gaps 21;

QY 9 YKNLYSDDIPHALREKRTTVAQLKQLOAETEPVIMFEDPETTRQMO 56
DB 77 YFNGYS-----LPERTEVTVTVERAKRETLLKTDDELKFKELLHLMSEEPKSGCKLE 130

QY 57 STRDGM-----LFDYLDKXHGFRQBYLDTLYRYAKFQYECNYSGAAYLYF 104
DB 131 KKDLDLNMPPDMKSGKINDYSDP-----KEYISKYDYALFSIH-----TPLAYF 179

QY 105 FR---VLVATDRNALSSLMGKLASEILMONWDAAMEDLTRKKTIDNNSV----SSPLQ 157
DB 180 VKSNLVRKTKCTKYGSDSYKIAYQAMLQKF---LLSIVQFKDRHNRLLLEPFSSPIA 236

QY 158 SLQORTWLIHWSLFVFNHPKGRDNIIDLFLYQPOLNAIOTMCPHILRYLTAVITNKD 217

```
Db 237 DEKRKNCL---TKFVIQDENKNSSTIADLCVWLKSREIKLQ-----ILLLEIIGLNDLD 288
Qy 218 -----VRKRQVL--KDLVKVLIQOESYTYKDPITBEFVECLYVNFDFGQAKX 262
Db 289 WNFPRDEKKYKLLKKRSLMLTKGLVRRSRKKTSEK---KGIERITTSLDY----- 339
Qy 263 LRECESVLNDFEL-VACL-----EDFIENARLFIFETFCRIHQICISINMLADK 310
Db 340 ---CEQL---DLYLDRACILDLILLSSETPNPDAIEASNGTIQE-----HK 378
Qy 311 LNMTPBEAERWIVNLIRNARLDKIDSKLGHVYVGNNAVSPYQOVIETKTSLSFRSQMLA 370
Db 379 KNILDKSKEASLVGFI-NYVLIFPYFNKKVPHAV-----EFIIQKLKGPSMRPKRAL 428
Qy 371 MNIEKKLNQNSSEAPN 387
Db 429 -----KKVNDSTNVSSPN 441
```

RESULT 15

```
JC7595
scavenger receptor with C-type lectin type I - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C:Accession: JC7595
R:Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001
A:Title: Molecular cloning and functional characterization of a human scavenger receptor
A:Reference number: JC7595; PMID:21092718; PMID:11162630
A:Contents: Placenta
A:Accession: JC7595
A:Molecule type: mRNA
A:Residues: 1-742 <NAK>
A:Cross-references: UNIPROT:Q9BYH7; DDBJ:AB038518
C:Comment: This receptor, a member of the scavenger receptor family, belonging to the tyrosine kinase family, is an important role in host defense. It forms a timer and plays a role in recognizing infected cells.
C:Genetics:
A:Gene: srcl-I
A:Map position: 18p11.32
C:Keywords: coiled coil; glycoprotein; transmembrane protein
F:1-39/Domain: cytosolic (amino-terminus) #status predicted <CYT>
F:16-19/Region: internalization signal YKRF
F:40-56/Domain: transmembrane #status predicted <TM>
F:57-112/Domain: extracellular #status predicted <EXT>
F:113-335/Domain: coiled coil #status predicted <COC>
F:369-384/Region: serine/threonine-rich #status predicted <SRC>
F:443-589/Domain: collagen-like #status predicted <COL>
F:607-732/Domain: C-type lectin/carbohydrate recognition domain #status predicted <CRD>
```

```
Query Match 5.0%; Score 103.5; DB 2; Length 742;
Best Local Similarity 21.2%; Pred. No. 8;
Matches 76; Conservative 52; Mismatches 115; Indels 115; Gaps 14;

Qy 6 MDVYKLYSDIPHALREKRTTVVAOLKQLOAETEPVKMFEDPETTRQMSQTRDGRMLF 65
Db 72 METSRQTYDD-----KLTAVESDLKLGDT-----GKKAI 102

Qy 66 DYLAHKGHFRQYELDTLYRKYAQYECGNYSGAAYLYFFRVLPVATDRNALSSWGKLA 125
Db 103 STNSELSTFRSDILDLRQ-----LREITEKTSKN-----KDT 135

Qy 126 SEILMQNWDAMEDLTRLKETIDNNS-----VSSPLQ-----SLQORTWLIHWSL-FV 172
Db 136 LEKLGASGALVDROSQLKTELENSFLITTVNKTQAVNGYVTVNLQDQTSVLQGNLQ 195

Qy 173 FENHPKGRDNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDKVRKQVLDLVKVI 232
Db 196 MYSHNVVIMNINLNITVQOQRN-----LITN-----LQRSVDDTSQAI 234

Qy 233 QOESYTYKDPITBEFVECLYVNFDFGQAKKRECESVLVNDFELVACLEDIE--NARLF 290
Db 235 QRIKNDQNLQVYF---LQAKKDTWLKEKVSQSLQTLAANNSALAKANNDTLEDNNSQLN 291
```

```
Qy 291 IF-----ETFCRIHQICISINM-----LADKLNMTPEAERW---IVNLIEN 328
Db 292 SFTGQMNITTTISQANEQNKLQDLHLKDAENRTAIKFQLEERFQLFETDIVNIISN 349
```

Search completed: March 8, 2005, 20:08:09
Job time : 50 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:00:57 ; Search time 43 Seconds
(without alignments)
687.466 Million cell updates/sec

Title: US-10-783-415-4
Perfect score: 2071
Sequence: 1 MVDFAWDYKLYNLYDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table: BLOSUM62
Gap0 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2071	100.0	396	3 US-08-875-847B-2	Sequence 2, Appli
2	2071	100.0	396	3 US-08-875-847B-4	Sequence 4, Appli
3	2071	100.0	396	3 US-09-378-842-2	Sequence 2, Appli
4	2071	100.0	396	3 US-09-378-842-4	Sequence 4, Appli
5	2071	100.0	396	4 US-09-858-152B-2	Sequence 2, Appli
6	2071	100.0	396	4 US-09-858-152B-4	Sequence 4, Appli
7	2071	100.0	445	4 US-09-976-594-841	Sequence 841, App
8	106	5.1	1575	4 US-09-917-254-83	Sequence 83, Appl
9	106	5.1	1575	4 US-09-949-016-6743	Sequence 6743, Ap
10	104	5.0	489	4 US-09-248-796A-19030	Sequence 19030, A
11	101.5	4.9	313	4 US-09-248-796A-23197	Sequence 23197, A
12	101.5	4.9	578	1 US-08-766-014-4	Sequence 4, Appli
13	101.5	4.9	608	1 US-08-766-014-3	Sequence 3, Appli
14	100.5	4.9	1196	4 US-09-107-532A-3944	Sequence 3944, Ap
15	100	4.8	1559	4 US-09-949-016-10190	Sequence 10190, A
16	98.5	4.8	872	1 US-08-766-014-2	Sequence 2, Appli
17	98.5	4.8	1056	4 US-09-079-030-217	Sequence 217, App
18	98.5	4.8	2663	4 US-09-538-092-1252	Sequence 1252, Ap
19	98.5	4.8	4536	4 US-09-180-422B-27	Sequence 27, Appl
20	98.5	4.8	4536	4 US-09-079-030-1	Sequence 1, Appli
21	98.5	4.8	4563	4 US-09-538-092-842	Sequence 842, App
22	98	4.7	2662	4 US-09-595-684B-31	Sequence 31, Appl
23	97.5	4.7	1979	4 US-09-949-016-6468	Sequence 6468, Ap
24	97.5	4.7	2047	4 US-09-949-016-7404	Sequence 7404, Ap
25	97.5	4.7	4563	4 US-09-108-006C-1	Sequence 1, Appli
26	96.5	4.7	2807	4 US-09-543-681A-4980	Sequence 4980, Ap
27	95	4.6	1011	4 US-09-489-039A-11808	Sequence 11808, A

28	94.5	4.6	789	4 US-09-248-796A-19294	Sequence 19294, A
29	94.5	4.6	913	3 US-08-971-089-4	Sequence 4, Appli
30	94	4.5	738	4 US-09-107-532A-6218	Sequence 6218, Ap
31	93.5	4.5	721	4 US-08-851-435-2	Sequence 2, Appli
32	93.5	4.5	746	4 US-08-851-435-6	Sequence 6, Appli
33	93	4.5	420	4 US-09-270-767-42980	Sequence 42980, A
34	93	4.5	599	2 US-08-910-551B-2	Sequence 2, Appli
35	93	4.5	916	4 US-09-688-078-9	Sequence 9, Appli
36	92.5	4.5	390	4 US-09-107-532A-4359	Sequence 4359, Ap
37	92.5	4.5	530	4 US-09-270-767-41545	Sequence 41545, A
38	92.5	4.5	792	4 US-09-543-681A-6617	Sequence 6617, Ap
39	92	4.4	660	4 US-09-248-796A-14578	Sequence 14578, A
40	92	4.4	991	4 US-09-248-796A-15239	Sequence 15239, A
41	91.5	4.4	572	4 US-09-134-000C-4237	Sequence 4237, Ap
42	90.5	4.4	568	3 US-09-134-001C-3768	Sequence 3768, Ap
43	90.5	4.4	1753	4 US-09-248-796A-19154	Sequence 19154, A
44	90	4.3	655	4 US-09-949-016-6759	Sequence 6759, Ap
45	90	4.3	669	4 US-09-949-016-10241	Sequence 10241, A

ALIGNMENTS

RESULT 1
US-08-875-847B-2
; Sequence 2, Application US/08875847B
; Patent No. 6255105
; GENERAL INFORMATION:
; APPLICANT: The Government of the United
; APPLICANT: States of America as represented by the
; APPLICANT: Secretary, Department of Health and Human
; APPLICANT: Services, Callahan, Robert; Marchetti,
; APPLICANT: Antonio; Buticita, Fiamma; Smith, Gilbert H.
; TITLE OF INVENTION: Nucleotide And Deduced
; TITLE OF INVENTION: Amino Acid Sequences Of A New Tumor Gene,
; TITLE OF INVENTION: Int6, And The Use Of Reagents Derived From
; TITLE OF INVENTION: These Sequences In Diagnostic Assays.
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,847B
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/385,998
; FILING DATE: 09-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feller
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4179PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-6840
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown


```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,842
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,847
; FILING DATE: 09-FEB-1996
; APPLICATION NUMBER: 08/385,998
; FILING DATE: 09-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4179PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
;
US-09-378-842-2

Query Match 100.0%; Score 2071; DB 3; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.8e-212;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDYKKNLYSDDI PHALREKRTTVVAQLKQLQAETEPIVKMFEDPETTRQMOSTRD 60
DB 1 MVDFAVDYKKNLYSDDI PHALREKRTTVVAQLKQLQAETEPIVKMFEDPETTRQMOSTRD 60
QY 61 GRMLFDYLDKXGFRQEYLDTRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYLDKXGFRQEYLDTRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAMEDLTRKKTIDNNSSVSPQLSQRTWLIIHWSLFFVFNHPKGR 180
DB 121 WGLKASEILMQNWDAAMEDLTRKKTIDNNSSVSPQLSQRTWLIIHWSLFFVFNHPKGR 180
QY 181 DNIIDLFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
DB 181 DNIIDLFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
QY 241 DPITEFVECLYVNFDPDGAOKKLRECESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300
DB 241 DPITEFVECLYVNFDPDGAOKKLRECESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARLDKIDSKLGHVVMGNNVSPYQQVIEKTK 360
DB 301 CISINMLADKLNMTPEEAERWIVNLRNARLDKIDSKLGHVVMGNNVSPYQQVIEKTK 360
QY 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
DB 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 4
US-09-378-842-4
; Sequence 4, Application US/09378842
; Patent No. 6342392
; GENERAL INFORMATION:
; APPLICANT: The Government of the United
; APPLICANT: States of America as represented by the
; APPLICANT: Secretary, Department of Health and Human
; APPLICANT: Services; Callahan, Robert; Marchetti,
; APPLICANT: Antonio; Buttitta, Fiamma; Smith, Gilbert H.
; TITLE OF INVENTION: Nucleotide And Deduced
; TITLE OF INVENTION: Amino Acid Sequences Of A New Tumor Gene,
; TITLE OF INVENTION: Int6, And The Use Of Reagents Derived From
```

```

; TITLE OF INVENTION: These Sequences In Diagnostic Assays,
; TITLE OF INVENTION: Vaccines, Immunotherapy And Gene Therapy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,842
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,847
; FILING DATE: 09-FEB-1996
; APPLICATION NUMBER: 08/385,998
; FILING DATE: 09-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4179PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
;
US-09-378-842-4

Query Match 100.0%; Score 2071; DB 3; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.8e-212;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDYKKNLYSDDI PHALREKRTTVVAQLKQLQAETEPIVKMFEDPETTRQMOSTRD 60
DB 1 MVDFAVDYKKNLYSDDI PHALREKRTTVVAQLKQLQAETEPIVKMFEDPETTRQMOSTRD 60
QY 61 GRMLFDYLDKXGFRQEYLDTRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYLDKXGFRQEYLDTRYAKFOYECGNTSGAAEYLYFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAMEDLTRKKTIDNNSSVSPQLSQRTWLIIHWSLFFVFNHPKGR 180
DB 121 WGLKASEILMQNWDAAMEDLTRKKTIDNNSSVSPQLSQRTWLIIHWSLFFVFNHPKGR 180
QY 181 DNIIDLFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
DB 181 DNIIDLFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
QY 241 DPITEFVECLYVNFDPDGAOKKLRECESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300
DB 241 DPITEFVECLYVNFDPDGAOKKLRECESVLVNDFFLVACLEDFIENARLFIPTFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARLDKIDSKLGHVVMGNNVSPYQQVIEKTK 360
DB 301 CISINMLADKLNMTPEEAERWIVNLRNARLDKIDSKLGHVVMGNNVSPYQQVIEKTK 360
QY 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
DB 361 SLFSRQMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
```

```
RESULT 5
US-09-858-152B-2
; Sequence 2, Application US/09858152B
; Patent No. 6737251
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Fiamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858,152B
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 09/858,152
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Murine INT6
US-09-858-152B-2

Query Match 100.0%; Score 2071; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.8e-212;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDYKLYSDDI PHALREKRTTVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60
Db 1 MVDFAVDYKLYSDDI PHALREKRTTVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60
Qy 61 GRMLFDYLDKXGFRQEVLDLYRYAKFOYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
Db 61 GRMLFDYLDKXGFRQEVLDLYRYAKFOYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
Qy 121 WGLASEILMQNWDAAEDLTRLKETIDNNSSVSPLOSLQORTWLIHWSLVFFFNHPKGR 180
Db 121 WGLASEILMQNWDAAEDLTRLKETIDNNSSVSPLOSLQORTWLIHWSLVFFFNHPKGR 180
Qy 181 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVKRQVQLKOLVKVIQOESYTYK 240
Db 181 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVKRQVQLKOLVKVIQOESYTYK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVLVACLEDFIENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVLVACLEDFIENARLFIETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLSFRSOMLANNIEKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSFRSOMLANNIEKLNQNSRSEAPNATQDSGFY 396

RESULT 6
US-09-858-152B-4
; Sequence 4, Application US/09858152B
; Patent No. 6737251
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Fiamma
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
```

```
; CURRENT APPLICATION NUMBER: US/09/858,152B
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 09/858,152
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-152B-4

Query Match 100.0%; Score 2071; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.8e-212;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDYKLYSDDI PHALREKRTTVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60
Db 1 MVDFAVDYKLYSDDI PHALREKRTTVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60
Qy 61 GRMLFDYLDKXGFRQEVLDLYRYAKFOYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
Db 61 GRMLFDYLDKXGFRQEVLDLYRYAKFOYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
Qy 121 WGLASEILMQNWDAAEDLTRLKETIDNNSSVSPLOSLQORTWLIHWSLVFFFNHPKGR 180
Db 121 WGLASEILMQNWDAAEDLTRLKETIDNNSSVSPLOSLQORTWLIHWSLVFFFNHPKGR 180
Qy 181 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVKRQVQLKOLVKVIQOESYTYK 240
Db 181 DNIIDFLYQPOLYNAIQTMCPHILRYLTAVITNKDVKRQVQLKOLVKVIQOESYTYK 240
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVLVACLEDFIENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVLVACLEDFIENARLFIETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLSFRSOMLANNIEKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSFRSOMLANNIEKLNQNSRSEAPNATQDSGFY 396

RESULT 7
US-09-976-594-841
; Sequence 841, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 841
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 1330056CD1
US-09-976-594-841

Query Match 100.0%; Score 2071; DB 4; Length 445;
Best Local Similarity 100.0%; Pred. No. 3.4e-212;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MVDFAVDVYKNLYSDDI PHALREKRTTVVAQLKQLQAEPIVKMPEDPETTRQMSQTRD 60
DQ 50 MVDFAVDVYKNLYSDDI PHALREKRTTVVAQLKQLQAEPIVKMPEDPETTRQMSQTRD 109
QY 61 GRMLFDYLDKGFQREYLDITLYRYAKFOYECGNTSGAAEYLYFFRVLVDPATDRNALSSL 120
DQ 110 GRMLFDYLDKGFQREYLDITLYRYAKFOYECGNTSGAAEYLYFFRVLVDPATDRNALSSL 169
QY 121 WGLKASEILMOWNDAMEDLTRKETIDNNSVSSPLOSQOQRTWLHLSLFFVFNHPKGR 180
DQ 170 WGLKASEILMOWNDAMEDLTRKETIDNNSVSSPLOSQOQRTWLHLSLFFVFNHPKGR 229
QY 181 DNIIDFLYQPOLYNAITQMCPHILRYLTAVITNKDVRKRQVLKDLVKVYQOESYTYK 240
DQ 230 DNIIDFLYQPOLYNAITQMCPHILRYLTAVITNKDVRKRQVLKDLVKVYQOESYTYK 289
QY 241 DPITFVECLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIETFCRIHQ 300
DQ 290 DPITFVECLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIETFCRIHQ 349
QY 301 CISINMLADKLNMTPEEAERWILNIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
DQ 350 CISINMLADKLNMTPEEAERWILNIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
DQ 410 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 8

US-09-917-254-83
; Sequence 83, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:

; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224 (JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 83
; LENGTH: 1575
; TYPE: PRT
; ORGANISM: Homo Sapiens

US-09-917-254-83

Query Match 5.1%; Score 106; DB 4; Length 1575;
Best Local Similarity 21.8%; Pred. No. 0.16;
Matches 86; Conservative 49; Mismatches 133; Indels 126; Gaps 19;

QY 48 DPETTRQMSQTRDGRMLFDYLDKGFQREYLDITLYRYAKFOYE-----CGNYSG 97
DQ 318 DPENT-----LLALKKPEAQ--LPVYPPAAAMYQNELFNQKQNTWNYLA 361
QY 98 -----AAEYLYFFRVLVDPATDRNALSSLGKGLASEIL-MQWDAAMEDLTRKETIDNN 150
DQ 362 HEELIAVEMLSAVALNQALNESDLVSQNLRSIPAIGLNLDKA-----YVERYANT 415
QY 151 SVSSPLOSQOQRTWLHLSLFFVFNHPKGRDNIIDFLYQPO-----YLN-AIQT 199
DQ 416 LLSVKLEVLSSQODNLSWNEI-----QNCIDMVNAQIQEENDRVVAVGYINEAIDE 466
QY 200 MCPHILRYLTAVITNKDVRKRQVLKDLVKVYQOESYTYKDPITFVECLYVNFDFGA 259
DQ 467 GNP--LRLETLLTLLPTANIS-----DVPDPAHAQHYQDVL-----YHAK 502
QY 260 OKKLRECESVLNDFLVACLEDFIENARLFIETFCRIHQICISINMLADKLNMTPEEA 319
DQ 503 SQKLGDSSEV-----SKVLWDE-----IQQAVDEANVDEDRAK 536

RESULT 10

US-09-248-796A-19030
; Sequence 19030, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

US-09-248-796A-19030

QY 320 RWIVNLIR-NARLDAKIDSKLGHVVMGN-----NAVSP-----YQOVIETKSL-----SF 364
DQ 537 QWTVLVVDVNOCLGKSSDILSVLKSSSTNSANDIIPEDAKYDYDALVKAKELKSERVSS 596
QY 365 RSQMLAMNIEKKLN--QNSRSEAPNWTQDSGFY 396
DQ 597 DGSMLKLNHLKKYDYNYNTDSKSSSWVTPESCFY 630

QY 320 RWIVNLIR-NARLDAKIDSKLGHVVMGN-----NAVSP-----YQOVIETKSL-----SF 364
DQ 537 QWTVLVVDVNOCLGKSSDILSVLKSSSTNSANDIIPEDAKYDYDALVKAKELKSERVSS 596
QY 365 RSQMLAMNIEKKLN--QNSRSEAPNWTQDSGFY 396
DQ 597 DGSMLKLNHLKKYDYNYNTDSKSSSWVTPESCFY 630

RESULT 9

US-09-949-016-6743
; Sequence 6743, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6743
; LENGTH: 1575
; TYPE: PRT
; ORGANISM: Human

US-09-949-016-6743

Query Match 5.1%; Score 106; DB 4; Length 1575;
Best Local Similarity 21.8%; Pred. No. 0.16;
Matches 86; Conservative 49; Mismatches 133; Indels 126; Gaps 19;

QY 48 DPETTRQMSQTRDGRMLFDYLDKGFQREYLDITLYRYAKFOYE-----CGNYSG 97
DQ 318 DPENT-----LLALKKPEAQ--LPVYPPAAAMYQNELFNQKQNTWNYLA 361
QY 98 -----AAEYLYFFRVLVDPATDRNALSSLGKGLASEIL-MQWDAAMEDLTRKETIDNN 150
DQ 362 HEELIAVEMLSAVALNQALNESDLVSQNLRSIPAIGLNLDKA-----YVERYANT 415
QY 151 SVSSPLOSQOQRTWLHLSLFFVFNHPKGRDNIIDFLYQPO-----YLN-AIQT 199
DQ 416 LLSVKLEVLSSQODNLSWNEI-----QNCIDMVNAQIQEENDRVVAVGYINEAIDE 466
QY 200 MCPHILRYLTAVITNKDVRKRQVLKDLVKVYQOESYTYKDPITFVECLYVNFDFGA 259
DQ 467 GNP--LRLETLLTLLPTANIS-----DVPDPAHAQHYQDVL-----YHAK 502
QY 260 OKKLRECESVLNDFLVACLEDFIENARLFIETFCRIHQICISINMLADKLNMTPEEA 319
DQ 503 SQKLGDSSEV-----SKVLWDE-----IQQAVDEANVDEDRAK 536
QY 320 RWIVNLIR-NARLDAKIDSKLGHVVMGN-----NAVSP-----YQOVIETKSL-----SF 364
DQ 537 QWTVLVVDVNOCLGKSSDILSVLKSSSTNSANDIIPEDAKYDYDALVKAKELKSERVSS 596
QY 365 RSQMLAMNIEKKLN--QNSRSEAPNWTQDSGFY 396
DQ 597 DGSMLKLNHLKKYDYNYNTDSKSSSWVTPESCFY 630

```

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19030
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-19030

```

Query Match 5.0%; Score 104; DB 4; Length 489;
Best Local Similarity 20.5%; Pred. No. 0.04;
Matches 75; Conservative 68; Mismatches 139; Indels 84; Gaps 16;

[illegible]

```

RESULT 11
US-09-248-796A-23197
; Sequence 23197, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 23197
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-23197

```

Query Match 4.9%; Score 101.5; DB 4; Length 313;

[illegible]

RESULT 12

US-08-766-014-4
; Sequence 4, Application US/08766014
; Patent No. 5744312
; GENERAL INFORMATION:
; APPLICANT: Mamone, Joseph A.
; APPLICANT: Davis, Maria
; APPLICANT: Sha, Dan
; TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,014
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 60/008,688
; FILING DATE: December 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 223/104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-766-014-4

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.


OM protein - protein search, using sw model

Run on: March 8, 2005, 20:08:59 ; Search time 44 Seconds
(without alignments)
671.842 Million cell updates/sec

Title: US-10-783-415-4

Perfect score: 396

Sequence: 1 MVDPMVYKLYSDDIPHA.....LNQNSRSEAPNWTQDSGFY 396

Scoring table:  Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	100.0	396	3	US-08-875-847B-2
2	396	100.0	396	3	US-08-875-847B-4
3	396	100.0	396	3	US-08-378-842-2
4	396	100.0	396	3	US-08-378-842-4
5	396	100.0	396	4	US-09-858-152B-2
6	396	100.0	396	4	US-09-858-152B-4
7	396	100.0	445	4	US-09-976-594-841
8	7	1.8	53	4	US-09-270-767-46143
9	7	1.8	55	4	US-09-270-767-35620
10	7	1.8	55	4	US-09-270-767-50837
11	7	1.8	62	4	US-09-621-976-7296
12	7	1.8	70	4	US-09-621-976-4933
13	7	1.8	74	3	US-09-025-151-8
14	7	1.8	74	4	US-09-637-240-8
15	7	1.8	81	4	US-09-621-976-7295
16	7	1.8	81	4	US-09-621-976-7298
17	7	1.8	94	1	US-08-167-035-23
18	7	1.8	94	1	US-08-208-887A-23
19	7	1.8	94	2	US-08-539-005-23
20	7	1.8	94	3	US-08-280-598-25
21	7	1.8	120	4	US-09-489-039A-10583
22	7	1.8	133	4	US-09-034-418-6
23	7	1.8	148	4	US-09-489-039A-14211
24	7	1.8	172	4	US-09-538-092-450
25	7	1.8	175	4	US-09-248-796A-19585
26	7	1.8	182	2	US-08-467-603-89
27	7	1.8	182	2	US-08-466-793-89

28	7	1.8	182	2	US-08-491-861A-89	Sequence 89, Appl
29	7	1.8	182	4	US-09-374-871A-89	Sequence 89, Appl
30	7	1.8	196	4	US-09-270-767-36727	Sequence 36727, A
31	7	1.8	196	4	US-09-270-767-51944	Sequence 51944, A
32	7	1.8	207	3	US-09-302-769-36	Sequence 36, Appl
33	7	1.8	238	4	US-09-489-039A-9350	Sequence 9350, Ap
34	7	1.8	238	4	US-09-248-796A-16553	Sequence 16553, A
35	7	1.8	259	4	US-09-248-796A-20877	Sequence 20877, A
36	7	1.8	307	4	US-09-710-279-522	Sequence 522, App
37	7	1.8	308	3	US-09-134-001C-4400	Sequence 4400, Ap
38	7	1.8	314	3	US-08-653-648A-16	Sequence 16, Appl
39	7	1.8	314	4	US-09-564-418-7	Sequence 7, Appl
40	7	1.8	373	4	US-09-149-476-374	Sequence 374, App
41	7	1.8	375	1	US-08-027-986-1	Sequence 1, Appl
42	7	1.8	375	1	US-08-027-986-2	Sequence 2, Appl
43	7	1.8	403	4	US-08-887-534A-83	Sequence 83, Appl
44	7	1.8	403	4	US-09-527-431-83	Sequence 83, Appl
45	7	1.8	403	4	US-09-446-861-83	Sequence 83, Appl

ALIGNMENTS

RESULT 1
US-08-875-847B-2
; Sequence 2, Application US/08875847B
; Patent No. 6255105
; GENERAL INFORMATION:
; APPLICANT: The Government of the United
; APPLICANT: States of America as represented by the
; APPLICANT: Secretary, Department of Health and Human
; APPLICANT: Services; Callahan, Robert; Marchetti,
; APPLICANT: Antonio; Nuccitella, Fiamma; Smith, Gilbert H.
; TITLE OF INVENTION: Nucleotide And Deduced
; TITLE OF INVENTION: Amino Acid Sequences Of A New Tumor Gene,
; TITLE OF INVENTION: Int6, And The Use Of Reagents Derived From
; TITLE OF INVENTION: These Sequences In Diagnostic Assays,
; TITLE OF INVENTION: Vaccines, Immunotherapy And Gene Therapy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154

COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,847B
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/385,998
; FILING DATE: 09-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4179PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown


```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,842
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,847
; FILING DATE: 09-FEB-1996
; APPLICATION NUMBER: 08/385,998
; FILING DATE: 09-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4179PCT
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; US-09-378-842-2

```

```

Query Match      100.0%; Score 396; DB 3; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDVYKLYSDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPETTRQMQSTRD 60
Db 1 MVDFAVDVYKLYSDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPETTRQMQSTRD 60

Qy 61 GRMLFDYLADKHGFRQEVLDLTYRYAKFOYECGNYSGAAEYLYFVRVLVPATDRNALSSL 120
Db 61 GRMLFDYLADKHGFRQEVLDLTYRYAKFOYECGNYSGAAEYLYFVRVLVPATDRNALSSL 120

Qy 121 WGLASEILMNQWDAAMEDLTRKLETIDNNSSVSPLOSQLOQTWLIHWSLVFFFNHPKGR 180
Db 121 WGLASEILMNQWDAAMEDLTRKLETIDNNSSVSPLOSQLOQTWLIHWSLVFFFNHPKGR 180

Qy 181 DNIIDFLYQPYQYLNATQTCMCPHILRYLTAVITNKDVKRQVLRKQVLRKQVLRKQVLRKQV 240
Db 181 DNIIDFLYQPYQYLNATQTCMCPHILRYLTAVITNKDVKRQVLRKQVLRKQVLRKQVLRKQV 240

Qy 241 DPITEFVECLYVNFDFDGAQKKLRECEESLVNDFFLVACLEDFTENARLPIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKKLRECEESLVNDFFLVACLEDFTENARLPIFETFCRIHQ 300

Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360

Qy 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

```

```

RESULT 4
US-09-378-842-4
; Sequence 4, Application US/09378842
; Patent No. 6342392
; GENERAL INFORMATION:
; APPLICANT: The Government of the United
; APPLICANT: States of America as represented by the
; APPLICANT: Secretary, Department of Health and Human
; APPLICANT: Services; Callahan, Robert; Marchetti,
; APPLICANT: Antonio; Buttitta, Fiamma; Smith, Gilbert H.
; TITLE OF INVENTION: Nucleotide And Deduced
; TITLE OF INVENTION: Amino Acid Sequences Of A New Tumor Gene,
; TITLE OF INVENTION: Intc6, And The Use Of Reagents Derived From

```

```

; TITLE OF INVENTION: These Sequences In Diagnostic Assays,
; TITLE OF INVENTION: Vaccines, Immunotherapy And Gene Therapy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,842
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,847
; FILING DATE: 09-FEB-1996
; APPLICATION NUMBER: 08/385,998
; FILING DATE: 09-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4179PCT
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; US-09-378-842-4

```

```

Query Match      100.0%; Score 396; DB 3; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDVYKLYSDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPETTRQMQSTRD 60
Db 1 MVDFAVDVYKLYSDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPETTRQMQSTRD 60

Qy 61 GRMLFDYLADKHGFRQEVLDLTYRYAKFOYECGNYSGAAEYLYFVRVLVPATDRNALSSL 120
Db 61 GRMLFDYLADKHGFRQEVLDLTYRYAKFOYECGNYSGAAEYLYFVRVLVPATDRNALSSL 120

Qy 121 WGLASEILMNQWDAAMEDLTRKLETIDNNSSVSPLOSQLOQTWLIHWSLVFFFNHPKGR 180
Db 121 WGLASEILMNQWDAAMEDLTRKLETIDNNSSVSPLOSQLOQTWLIHWSLVFFFNHPKGR 180

Qy 181 DNIIDFLYQPYQYLNATQTCMCPHILRYLTAVITNKDVKRQVLRKQVLRKQVLRKQVLRKQV 240
Db 181 DNIIDFLYQPYQYLNATQTCMCPHILRYLTAVITNKDVKRQVLRKQVLRKQVLRKQVLRKQV 240

Qy 241 DPITEFVECLYVNFDFDGAQKKLRECEESLVNDFFLVACLEDFTENARLPIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKKLRECEESLVNDFFLVACLEDFTENARLPIFETFCRIHQ 300

Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQQVIEKTK 360

Qy 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSPRSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

```

```
RESULT 5
US-09-858-152B-2
; Sequence 2, Application US/09858152B
; Patent No. 6737251
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Fiama
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
; CURRENT APPLICATION NUMBER: US/09/858.152B
; PRIOR FILING DATE: 2001-05-14
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Murine INT6
US-09-858-152B-2

Query Match 100.0%; Score 396; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDPMFVYKLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60
Db 1 MVDPMFVYKLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60
Qy 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
Db 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
Qy 121 WGLASEILMQNWDAAEDLTRKETIDNNSVSPQLSQLOQRTWLIIHWSLFFVFNHPKGR 180
Db 121 WGLASEILMQNWDAAEDLTRKETIDNNSVSPQLSQLOQRTWLIIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQPYLNAIQTCWPHILRYLTAVITNKDVKRRQVLDKLVKVIQOESYTK 240
Db 181 DNIIDFLYQPYLNAIQTCWPHILRYLTAVITNKDVKRRQVLDKLVKVIQOESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAOKLRECECVLNDFFLVACLEDFTENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAOKLRECECVLNDFFLVACLEDFTENARLFIETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLFSRQMLAMNTEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNTEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 6
US-09-858-152B-4
; Sequence 4, Application US/09858152B
; Patent No. 6737251
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Marchetti, Antonio
; APPLICANT: Buttitta, Fiama
; APPLICANT: Smith, Gilbert H.
; APPLICANT: Callahan, Robert
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF TUMOR GENE INT6
; FILE REFERENCE: 4239-59122
```

```
; CURRENT APPLICATION NUMBER: US/09/858.152B
; CURRENT FILING DATE: 2001-05-14
; PRIOR FILING DATE: 09/858.152
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 4
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-152B-4

Query Match 100.0%; Score 396; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDPMFVYKLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60
Db 1 MVDPMFVYKLYSDDDIPHALREKRTTVAQLKQLAETEPVVKMFEDPETTRQMSTRD 60
Qy 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
Db 61 GRMLFDYLDKKGFRQBYLDTLYRYAKFQYECGNYSGAAEYLYFFRVLVPATDRNALSSL 120
Qy 121 WGLASEILMQNWDAAEDLTRKETIDNNSVSPQLSQLOQRTWLIIHWSLFFVFNHPKGR 180
Db 121 WGLASEILMQNWDAAEDLTRKETIDNNSVSPQLSQLOQRTWLIIHWSLFFVFNHPKGR 180
Qy 181 DNIIDFLYQPYLNAIQTCWPHILRYLTAVITNKDVKRRQVLDKLVKVIQOESYTK 240
Db 181 DNIIDFLYQPYLNAIQTCWPHILRYLTAVITNKDVKRRQVLDKLVKVIQOESYTK 240
Qy 241 DPITEFVECLYVNFDFDGAOKLRECECVLNDFFLVACLEDFTENARLFIETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAOKLRECECVLNDFFLVACLEDFTENARLFIETFCRIHQ 300
Qy 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Qy 361 SLFSRQMLAMNTEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLFSRQMLAMNTEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 7
US-09-976-594-841
; Sequence 841, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976.594
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 60/240.409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 841
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 1330056CD1
US-09-976-594-841

Query Match 100.0%; Score 396; DB 4; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 MVDPMVYKLYSDDIPHALREKRTTVAQLKOLQAEETPIVKMPEDPETTRMQSTRD 60
Db 50 MVDPMVYKLYSDDIPHALREKRTTVAQLKOLQAEETPIVKMPEDPETTRMQSTRD 109
Qy 61 GRMLFDYLDKKGFRQBYDLTYRYAKFQYECGNYSGAAEYLYFVRVLVPATDRLSSSL 120
Db 110 GRMLFDYLDKKGFRQBYDLTYRYAKFQYECGNYSGAAEYLYFVRVLVPATDRLSSSL 169
Qy 121 WGLKASILQNDWDAAMEDLTRKETIDNNSVSPLOSQLOORTWLIHWSLFPVFNHPKGR 180
Db 170 WGLKASILQNDWDAAMEDLTRKETIDNNSVSPLOSQLOORTWLIHWSLFPVFNHPKGR 229
Qy 181 DNIIDLFLYQOYLNAIQTCMCPHLRYLTAVITNKDKVRKQVLDLVKVIQOESYTYK 240
Db 230 DNIIDLFLYQOYLNAIQTCMCPHLRYLTAVITNKDKVRKQVLDLVKVIQOESYTYK 289
Qy 241 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVACLEDFTIENARLFIFETFCRIHQ 300
Db 290 DPITEFVECLYVNFDFDGAQKLRCECVLVNDFLVACLEDFTIENARLFIFETFCRIHQ 349
Qy 301 CISHNLADKLNMTPEAEERWVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 350 CISHNLADKLNMTPEAEERWVNLIRNARLDKIDSKLGHVVMGNNAVSPYQOVIETK 409
Qy 361 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 396
Db 410 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 445

RESULT 8

US-09-270-767-46143

; Sequence 46143, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 46143

; LENGTH: 53

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; US-09-270-767-46143

Query Match 1.8%; Score 7; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 QSTRDGR 62

Db 39 QSTRDGR 45

RESULT 9

US-09-270-767-35620

; Sequence 35620, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 35620

; LENGTH: 55

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

;

; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-35620

Query Match 1.8%; Score 7; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 ALSSSLWG 122

Db 47 ALSSSLWG 53

RESULT 10

US-09-270-767-50837

; Sequence 50837, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 50837

; LENGTH: 55

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-50837

Query Match 1.8%; Score 7; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 ALSSSLWG 122

Db 47 ALSSSLWG 53

RESULT 11

US-09-621-976-7296

; Sequence 7296, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano J.V.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 7296

; LENGTH: 62

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: 62

; OTHER INFORMATION: Xaa = *, Gly

US-09-621-976-7296

Query Match 1.8%; Score 7; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 RNALSSSL 120

Db 7 RNALSSSL 13

```
RESULT 12
US-09-621-976-4933
; Sequence 4933, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4933
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4933

Query Match 1.8%; Score 7; DB 4; Length 70;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 RNALSSL 120
Db 7 RNALSSL 13

RESULT 13
US-09-025-151-8
; Sequence 8, Application US/09025151
; Patent No. 6187535
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Fromont, Micheline
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: FAST AND EXHAUSTIVE METHOD FOR SELECTING A PREY
; TITLE OF INVENTION: POLYPEPTIDE INTERACTING WITH A BAIT POLYPEPTIDE OF
; TITLE OF INVENTION: INTEREST: APPLICATION TO THE CONSTRUCTION OF MAPS OF
; TITLE OF INVENTION: INTERACTORS POLYPEPTIDES
; FILE REFERENCE: 03495-0164
; CURRENT APPLICATION NUMBER: US/09/025,151
; CURRENT FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-025-151-8

Query Match 1.8%; Score 7; DB 3; Length 74;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 RDGRMLF 65
Db 14 RDGRMLF 20

RESULT 14
US-09-637-240-8
; Sequence 8, Application US/09637240
; Patent No. 6531284
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Fromont, Micheline
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: FAST AND EXHAUSTIVE METHOD FOR SELECTING A PREY
; TITLE OF INVENTION: POLYPEPTIDE INTERACTING WITH A BAIT POLYPEPTIDE OF
; TITLE OF INVENTION: INTEREST: APPLICATION TO THE CONSTRUCTION OF MAPS OF
; TITLE OF INVENTION: INTERACTORS POLYPEPTIDES
; FILE REFERENCE: 03495-0164
; CURRENT APPLICATION NUMBER: US/09/637,240
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/025,151
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-637-240-8

Query Match 1.8%; Score 7; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 RDGRMLF 65
Db 14 RDGRMLF 20

RESULT 15
US-09-621-976-7295
; Sequence 7295, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7295
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7295

Query Match 1.8%; Score 7; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 RNALSSL 120
Db 7 RNALSSL 13

Search completed: March 8, 2005, 20:23:34
Job time : 45 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 19:47:06 ; Search time 164 Seconds
(without alignments)
933.885 Million cell updates/sec

Title: US-10-783-415-4

Perfect score: 2071

Sequence: 1 MVDPMYVKNLYSDIDPHA.....LNQSRSEAPNWTQSGFY 396

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A: Geneseq 16Dec04: *

1: Geneseqp1980s: *

2: Geneseqp1990s: *

3: Geneseqp2000s: *

4: Geneseqp2001s: *

5: Geneseqp2002s: *

6: Geneseqp2003as: *

7: Geneseqp2003bs: *

8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2071	100.0	396	2 AAW02113	Aaw02113 Murine In
2	2071	100.0	396	2 AAW02112	Aaw02112 Human hom
3	2071	100.0	396	2 AAB47920	Aab47920 Murine In
4	2071	100.0	396	2 AAB47921	Aab47921 Human Int
5	2071	100.0	396	6 ABU05167	Abu05167 Human exp
6	2071	100.0	396	6 ABU05169	Abu05169 Human exp
7	2071	100.0	396	6 ABU05158	Abu05158 Human exp
8	2071	100.0	396	7 ADJ70258	Adj70258 Human exp
9	2071	100.0	396	8 ADN30996	Adn30996 Human Int
10	2071	100.0	396	8 ADN30994	Adn30994 Murine In
11	2071	100.0	445	5 ABB57306	Abb57306 Mouse isc
12	2071	100.0	445	6 ABR39935	Abr39935 Human pro
13	2071	100.0	445	6 ABU05160	Abu05160 Human exp
14	2071	100.0	445	6 ABU05163	Abu05163 Human exp
15	2071	100.0	445	6 ABU05157	Abu05157 Human exp
16	2071	100.0	445	6 ABU05166	Abu05166 Human exp
17	2071	100.0	445	6 ABU05159	Abu05159 Human exp
18	2071	100.0	445	6 ABU05162	Abu05162 Human exp
19	2071	100.0	445	6 ABU05164	Abu05164 Human exp
20	2071	100.0	445	6 ABU05168	Abu05168 Human exp
21	2071	100.0	445	7 ABM85381	Abm85381 Human pro
22	2071	100.0	445	8 ADL13112	Adl13112 Human ste
23	2071	100.0	445	8 ABM81005	Abm81005 Tumour-as
24	2067	99.8	445	6 ABU05161	Abu05161 Human exp
25	2029	98.0	445	6 ABU05165	Abu05165 Human exp

26	1279.5	61.8	320	7	ABM85380	Abm85380 Mouse pro
27	1220.5	58.9	435	4	ABM84579	Abm84579 Drosophil
28	1104.5	53.3	302	4	ABG00533	Abg00533 Novel hum
29	1104.5	53.3	302	4	ABG21243	Abg21243 Novel hum
30	1090.5	52.7	427	3	AAG34208	Aag34208 Zea may
31	1090.5	52.7	436	3	AAG34207	Aag34207 Zea may
32	1090.5	52.7	467	3	AAG34206	Aag34206 Zea may
33	903	43.6	418	8	ADM57227	Adm57227 A thalian
34	838.5	40.5	1202	4	ABG21244	Abg21244 Novel hum
35	838.5	40.5	1202	7	ADF60389	Adf60389 Human con
36	738.5	35.7	155	4	ABG21242	Abg21242 Novel hum
37	564	27.2	117	5	ABP43106	Abp43106 Human ova
38	537.5	26.0	197	3	AAG16552	Aag16552 Arabidops
39	537.5	26.0	250	3	AAG16551	Aag16551 Arabidops
40	537.5	26.0	274	3	AAG16550	Aag16550 Arabidops
41	223	10.8	45	4	AAM22004	Aam22004 Peptide #
42	223	10.8	45	4	ABB44381	Abb44381 Peptide #
43	223	10.8	45	4	AAM38359	Aam38359 Peptide #
44	223	10.8	45	4	ABB27240	Abb27240 Protein #
45	223	10.8	45	4	AAM78125	Aam78125 Human bon

ALIGNMENTS

RESULT 1

AAW02113
ID AAW02113 standard; protein; 396 AA.

XX AAW02113;

DT 15-MAY-1997 (first entry)

DE Murine Int6 protein associated with MMTV integration and tumour growth.

KW MMTV; mouse mammary tumour virus; Int6; breast cancer; neoplasia;

KW diagnosis; treatment; immunotherapy; vaccine; probe; primer.

OS Mus musculus.

FH Key Location/Qualifiers

FT Misc-difference 27

FT /note= "potential cAMP/cGMP-dependent protein kinase phosphorylation site"

FT Misc-difference 51

FT /note= "potential protein kinase C phosphorylation site"

FT Misc-difference 57

FT /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 90

FT /note= "potential tyrosine kinase phosphorylation site"

FT Misc-difference 94

FT /note= "potential glycosylation site"

FT Misc-difference 112

FT /note= "potential protein kinase C phosphorylation site"

FT Misc-difference 149

FT /note= "potential glycosylation site"

FT Misc-difference 214

FT /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 238

FT /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 314

FT /note= "potential casein kinase II phosphorylation site"

FT Misc-difference 363

FT /note= "potential protein kinase C phosphorylation site"

FT Misc-difference 381

FT /note= "potential casein kinase II phosphorylation site"

XX WO9624672-A1.

XX 15-AUG-1996.

PP 09-FEB-1996; 96WO-US001884.

```

PR 09-FEB-1995; 95US-00385998.
XX
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
XX WI; 1996-384444/38.
XX DR N-PSDB; AAT36177.
XX
XX PT DNA encoding Int6 tumour associated protein - and use of reagents derived
XX from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
XX
XX PS Disclosure; Page 60-61; 93pp; English.
XX
XX CC AAW02113 is the murine Int6 protein. The Int6 gene is located at
XX chromosome 15 of a mouse genome. The Int6 gene is associated with MMTV
XX (mouse mammary tumour virus) integration into a host genome during
XX tumorigenesis. Primers and probes may be derived from the Int6 gene
XX sequence and used for detection in assays to diagnose MMTV infection, or
XX any other Int6 gene integration. Antibodies against the Int6 protein can
XX be used in the same way. The DNA and protein may also be used gene
XX therapy and vaccines
XX
XX SQ Sequence 396 AA;
Query Match 100.0%; Score 2071; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDVYKNLYSDDI PHALREKRTTVAQLKQLQAEITEPIVKMFEDPETTRQMQSTRD 60
DB 1 MVDFAVDVYKNLYSDDI PHALREKRTTVAQLKQLQAEITEPIVKMFEDPETTRQMQSTRD 60
QY 61 GRMLFDYLADKHGFRQEYLDLTLYRYAKFQYCGNYSGAAYLYFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYLADKHGFRQEYLDLTLYRYAKFQYCGNYSGAAYLYFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIHWSLFFVFNHPKGR 180
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
DB 181 DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLYNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESVLYNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARDKIDSGLGHVVMGNNAVSPYQQVIEKTK 360
DB 301 CISINMLADKLNMTPEEAERWIVNLIRNARDKIDSGLGHVVMGNNAVSPYQQVIEKTK 360
QY 361 SLSPRSQMLAMNIEKKLNQSRSEAPNATQDSGFY 396
DB 361 SLSPRSQMLAMNIEKKLNQSRSEAPNATQDSGFY 396
RESULT 2
AAW02112
ID AAW02112 standard; protein; 396 AA.
XX
XX AC AAW02112;
XX
XX DT 14-MAY-1997 (first entry)
XX
XX DE Human homologue of Int6 protein associated with MMTV integration.
XX
XX KW MMTV; mouse mammary tumour virus; Int6; breast cancer; neoplasia;
XX diagnosis; treatment; immunotherapy; vaccine; probe; primer.
XX
XX OS Homo sapiens.
XX

```

```

PN WO9624672-A1.
XX
XX PD 15-AUG-1996.
XX
XX PF 09-FEB-1996; 96WO-US001884.
XX
XX PR 09-FEB-1995; 95US-00385998.
XX
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
XX WI; 1996-384444/38.
XX DR N-PSDB; AAT36148.
XX
XX PT DNA encoding Int6 tumour associated protein - and use of reagents derived
XX from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
XX
XX PS Claim 3; Page 62-63; 93pp; English.
XX
XX CC AAW02112 represents the human homologue of murine Int6 protein. The Int6
XX gene is located at chromosome 15 of the mouse genome and is associated
XX with MMTV (mouse mammary tumour virus) integration into a host genome
XX during tumorigenesis. Primers and probes may be derived from the Int6
XX cDNA sequence and used for the detection of the Int6 gene. These can be
XX used in assays to diagnose MMTV infection, or any other Int6 gene
XX integration. Antibodies against the Int6 protein can be used in the same
XX way
XX
XX SQ Sequence 396 AA;
Query Match 100.0%; Score 2071; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDVYKNLYSDDI PHALREKRTTVAQLKQLQAEITEPIVKMFEDPETTRQMQSTRD 60
DB 1 MVDFAVDVYKNLYSDDI PHALREKRTTVAQLKQLQAEITEPIVKMFEDPETTRQMQSTRD 60
QY 61 GRMLFDYLADKHGFRQEYLDLTLYRYAKFQYCGNYSGAAYLYFFRVLVPATDRNALSSL 120
DB 61 GRMLFDYLADKHGFRQEYLDLTLYRYAKFQYCGNYSGAAYLYFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIHWSLFFVFNHPKGR 180
DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
DB 181 DNIIDFLYQPYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDKLVKVIQOESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLYNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB 241 DPITEFVECLYVNFDFDGAQKLRCEESVLYNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARDKIDSGLGHVVMGNNAVSPYQQVIEKTK 360
DB 301 CISINMLADKLNMTPEEAERWIVNLIRNARDKIDSGLGHVVMGNNAVSPYQQVIEKTK 360
QY 361 SLSPRSQMLAMNIEKKLNQSRSEAPNATQDSGFY 396
DB 361 SLSPRSQMLAMNIEKKLNQSRSEAPNATQDSGFY 396
RESULT 3
AAW02112
ID AAW02112 standard; protein; 396 AA.
XX
XX AC AAW02112;
XX
XX DT 21-MAY-2002 (first entry)
XX
XX DE Murine Int6.

```

XX Murine; human; Int6; integration site; deregulation; neoplasia;
 KW mouse mammary tumour virus; MMV; cancer; immunotherapy; gene therapy;
 KW prenatal screening; foetus; vaccine.
 XX Mus musculus.

XX OS
 XX PH Location/Qualifiers
 FT 27 /note= "Phosphorylation site for cAMP/cGMP-dependent
 FT Modified-site protein kinase"
 FT 51
 FT Modified-site
 FT 57 /note= "Phosphorylation site for protein kinase C"
 FT Modified-site
 FT 90 /note= "Phosphorylation site for casein kinase II"
 FT Modified-site
 FT 94 /note= "Phosphorylation site for tyrosine kinase"
 FT Modified-site
 FT /note= "Phosphorylation site for tyrosine kinase and
 FT glycosylation site"
 FT 112
 FT Modified-site
 FT /note= "Phosphorylation site for protein kinase C"
 FT 149
 FT Modified-site
 FT /note= "Phosphorylation site for tyrosine kinase and
 FT glycosylation site"
 FT 214
 FT Modified-site
 FT /note= "Phosphorylation site for casein kinase II"
 FT 238
 FT Modified-site
 FT /note= "Phosphorylation site for casein kinase II"
 FT 314
 FT Modified-site
 FT /note= "Phosphorylation site for casein kinase II"
 FT 363
 FT Modified-site
 FT /note= "Phosphorylation site for protein kinase C"
 FT 381
 FT Modified-site
 FT /note= "Phosphorylation site for casein kinase II"

XX US6342392-B1.
 XX PN
 XX PD 29-JAN-2002.
 XX PF 23-AUG-1999; 99US-00378842.
 XX PR 09-FEB-1995; 95US-00385998.
 XX PR 09-FEB-1996; 96WO-US001884.
 XX PR 25-SEP-1997; 97US-00875847.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Marchetti A, Buttitta F, Smith GH, Callahan R;
 XX WPI; 1996-384444/38.
 XX DR N-PSDB; AAI72498.
 XX DNA encoding Int6 tumour associated protein - and use of reagents derived
 PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
 XX Disclosure; Col 35-38; 45pp; English.

XX This sequence shows murine Int6. Int6 is an integration site for mouse
 CC mammary tumour virus (MMTV), which causes deregulation of expression of
 CC cellular genes adjacent to the site of MMTV integration in mammary
 CC tumours. The Int6 protein has been found to be highly conserved across
 CC species, with Drosophila Int6 being 60% identical to human/mouse Int6.
 CC This indicates that Int6 is serving a basic life function. The method of
 CC the invention comprises assaying a sample to detect a human Int6 nucleic
 CC acid sequence, or its fragment, by contacting the sample with a sequence
 CC of at least 15 consecutive nucleotides of human Int6 cDNA or a
 CC conservative variant of it, where a disrupted expression or loss of
 CC expression of the variant is associated with neoplasia. The method is
 CC useful for prenatal screening of a foetus or to pre-symptomatically
 CC screen a subject at risk of having cancer. Detecting mutations in the
 CC Int6 gene can provide diagnostic and prognostic information. The nucleic
 CC acids and proteins are useful in immunotherapy, gene therapy or as

CC vaccines for treating or preventing cancer. The nucleic acids are useful
 CC as probes for isolating homologues of Int6 gene or for detecting
 CC mutations in the Int6 gene
 XX SQ Sequence 396 AA;
 Query Match 100.0%; Score 2071; DB 2; Length 396;
 Best Local Similarity 100.0%; Pred. No. 7,2e-193;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYDFAMDVYKXLYSDDIPHALREKRTTVVAQLKQLQAEPIVKMPFEDPETTRQMSTRD 60
 DB 1 MYDFAMDVYKXLYSDDIPHALREKRTTVVAQLKQLQAEPIVKMPFEDPETTRQMSTRD 60
 QY 61 GRLMFDYLDADKRGFRQEYLDTLRYAKFYQECGNTSGAAEYLYFFRVLVPATDRNALSSL 120
 DB 61 GRLMFDYLDADKRGFRQEYLDTLRYAKFYQECGNTSGAAEYLYFFRVLVPATDRNALSSL 120
 QY 121 MGKLAASEILMQNWDAAEDLTRLKETIDNNSVSSPLQSLQQRRTWLIHWSLFFVFNHPKGR 180
 DB 121 MGKLAASEILMQNWDAAEDLTRLKETIDNNSVSSPLQSLQQRRTWLIHWSLFFVFNHPKGR 180
 QY 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
 DB 181 DNIIDFLYQPOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
 QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIETFCRIHQ 300
 DB 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIETFCRIHQ 300
 QY 301 CUSINMLADKLNMTPEEAERWIVNLI RNARLDAKIDSKLGHVVMGNNAVSPYQOQVIEKTK 360
 DB 301 CUSINMLADKLNMTPEEAERWIVNLI RNARLDAKIDSKLGHVVMGNNAVSPYQOQVIEKTK 360
 QY 361 SLSPFSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396
 DB 361 SLSPFSQMLAMNIEKKLNQNSRSEAPNATQDSGFY 396

RESULT 4
 AAB47921
 ID AAB47921 standard; protein; 396 AA.
 XX
 AC AAB47921;
 XX 21-MAY-2002 (first entry)
 XX Human Int6.
 XX Murine; human; Int6; integration site; deregulation; neoplasia;
 KW mouse mammary tumour virus; MMV; cancer; immunotherapy; gene therapy;
 KW prenatal screening; foetus; vaccine; chromosome 8q22-Q24.
 XX Homo sapiens.
 XX US6342392-B1.
 XX 29-JAN-2002.
 XX 23-AUG-1999; 99US-00378842.
 XX 09-FEB-1995; 95US-00385998.
 XX 09-FEB-1996; 96WO-US001884.
 XX 25-SEP-1997; 97US-00875847.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Marchetti A, Buttitta F, Smith GH, Callahan R;
 XX WPI; 1996-384444/38.
 XX DR N-PSDB; AAI72499.
 XX DNA encoding Int6 tumour associated protein - and use of reagents derived
 PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.

PT from them in cancer gene therapy, vaccines, diagnosis and immunotherapy.
 PS Disclosure; Col 39-44; 45pp; English.

XX This sequence shows human Int6. The human Int6 coding sequence was
 CC isolated by using murine Int6 sequences as probes. Human Int6 is
 CC organised into 13 exons as is murine Int6, and contains a CA-repeat in
 CC the 7th intron. Human Int6 has been localised to chromosome 8, more
 CC specifically to 8q22-q24. Int6 is an integration site for mouse mammary
 CC tumour virus (MMTV), which causes deregulation of expression of cellular
 CC genes adjacent to the site of MMTV integration in mammary tumours. The
 CC Int6 protein has been found to be highly conserved across species, with
 CC Drosophila Int6 being 60% identical to human/mouse Int6. This indicates
 CC that Int6 is serving a basic life function. The method of the invention
 CC comprises assaying a sample to detect a human Int6 nucleic acid sequence,
 CC or its fragment, by contacting the sample with a sequence of at least 15
 CC consecutive nucleotides of human Int6 cDNA or a conservative variant of
 CC it, where a disrupted expression or loss of expression of the variant is
 CC associated with neoplasia. The method is useful for prenatal screening of
 CC a foetus or to pre-symptomatically screen a subject at risk of having
 CC cancer. Detecting mutations in the Int6 gene can provide diagnostic and
 CC prognostic information. The nucleic acids and proteins are useful in
 CC immunotherapy, gene therapy or as vaccines for treating or preventing
 CC cancer. The nucleic acids are useful as probes for isolating homologues
 CC of Int6 gene or for detecting mutations in the Int6 gene

XX Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 2; Length 396;
 Best Local Similarity 100.0%; Pred. No. 7.2e-193;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAOLKQAEETPIVKMFEDPETTRQMSSTRD 60
 DB 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAOLKQAEETPIVKMFEDPETTRQMSSTRD 60
 QY 61 GRMLFDYLADKHGFRQYEDLTLYRYAKFYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
 DB 61 GRMLFDYLADKHGFRQYEDLTLYRYAKFYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
 QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPQSQLOQRTWLIIHWSLFFVFNHPKGR 180
 DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPQSQLOQRTWLIIHWSLFFVFNHPKGR 180
 QY 181 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKVKRQVLDKLVKVIQQESYTYK 240
 DB 181 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKVKRQVLDKLVKVIQQESYTYK 240
 QY 241 DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
 DB 241 DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
 QY 301 CISINMLADKLNMTPEERIVNLIIRNARDADKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
 DB 301 CISINMLADKLNMTPEERIVNLIIRNARDADKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
 QY 361 SLSPRSQMLAMNIEKKLNQNSSEAPNATQDSGFY 396
 DB 361 SLSPRSQMLAMNIEKKLNQNSSEAPNATQDSGFY 396

RESULT 5

ABU05167

ID ABU05167 standard; protein; 396 AA.

XX

AC ABU05167;

XX

DT 29-JAN-2003 (first entry)

XX

DE Human expressed protein tag (EPT) #1833.

XX

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;

XX protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 XX cytoskeletal proteins, receptors or transcription factors), useful for
 XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 XX leukemia.

XX Example 2; SEQ ID NO 1833; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
 XX fragment of a kinase, phosphatase, protease, protease inhibitor,
 XX transporter, cytoskeletal protein, receptor or transcription factor. The
 XX polypeptide is useful as an immunogenic composition for eliciting in a
 XX mammal an immunogenic response directed against any of the purified
 XX polypeptide. The purified polypeptide, or the antibody that binds to this
 XX polypeptide, is useful for treating cancer. The polypeptide is also
 XX useful for identifying compounds that binds to a naturally processed
 XX class I or class II MHC-binding polypeptide. The polypeptides and
 XX polynucleotides are particularly useful for treating or preventing
 XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 XX lymphoma or leukemia. These are also useful for screening agents for
 XX treating the above mentioned diseases. This sequence represents an
 XX expressed protein tag (EPT) isolated from human tissue for translational
 XX profiling. Note: This sequence does not appear in the printed
 XX specification but was obtained in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 6; Length 396;

Best Local Similarity 100.0%; Pred. No. 7.2e-193;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAOLKQAEETPIVKMFEDPETTRQMSSTRD 60

DB 1 MVDFAVDVYKNLYSDDIIPHALREKRTTVVAOLKQAEETPIVKMFEDPETTRQMSSTRD 60

QY 61 GRMLFDYLADKHGFRQYEDLTLYRYAKFYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120

DB 61 GRMLFDYLADKHGFRQYEDLTLYRYAKFYECGNYSGAAEYLYFPRVLVPATDRNALSSL 120

QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPQSQLOQRTWLIIHWSLFFVFNHPKGR 180

DB 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPQSQLOQRTWLIIHWSLFFVFNHPKGR 180

QY 181 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKVKRQVLDKLVKVIQQESYTYK 240

DB 181 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKVKRQVLDKLVKVIQQESYTYK 240

QY 241 DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300

DB 241 DPITEFVECLYVNFDFDGAQKKLRECESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300

QY 301 CISINMLADKLNMTPEERIVNLIIRNARDADKIDSKLGHVVMGNNAVSPYQQVIEKTK 360

DB 301 CISINMLADKLNMTPEERIVNLIIRNARDADKIDSKLGHVVMGNNAVSPYQQVIEKTK 360

QY 361 SLSPRSQMLAMNIEKKLNQNSSEAPNATQDSGFY 396

DB 361 SLSPRSQMLAMNIEKKLNQNSSEAPNATQDSGFY 396

Db 241 DPITEFVECLYNVDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
Db 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360
QY 361 SLSFRSOMLANNIEKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSOMLANNIEKLNQNSRSEAPNWTQDSGFY 396

RESULT 6

ID ABU05169
ABU05169 standard; protein; 396 AA.

XX AC ABU05169;

DT 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1835.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

OS WO200278524-A2.

PN 10-OCT-2002.

PD 28-MAR-2002; 2002WO-US009671.

PF 28-MAR-2001; 2001US-0279495P.

PR 21-MAY-2001; 2001US-0292544P.

PR 08-AUG-2001; 2001US-0310801P.

PR 01-OCT-2001; 2001US-0326370P.

PR 04-DEC-2001; 2001US-0336780P.

PR 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.

XX Example 2; SEQ ID NO 1835; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDFAMDVYKNLYSDDIIPHALREKTTVVAAQLKQLAETEPYVKMFEDPETTRMQMSTRD 60
Db 1 MYDFAMDVYKNLYSDDIIPHALREKTTVVAAQLKQLAETEPYVKMFEDPETTRMQMSTRD 60

QY 61 GRMLFDYLADKHGFRQEVLDITLYRYAKFOYECGNTSGAAEYLYFFRVLVLPATDRNALSSL 120

Db 61 GRMLFDYLADKHGFRQEVLDITLYRYAKFOYECGNTSGAAEYLYFFRVLVLPATDRNALSSL 120

QY 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQQRWLIIHWSLFFVFNHPKGR 180

Db 121 WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQQRWLIIHWSLFFVFNHPKGR 180

QY 181 DNIIDFLYPOYLNAIOTMCPHILRYLTAVITNKDVKRBOVLKDLVKVIOQESYTYK 240

Db 181 DNIIDFLYPOYLNAIOTMCPHILRYLTAVITNKDVKRBOVLKDLVKVIOQESYTYK 240

QY 241 DPITEFVECLYNVDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300

Db 241 DPITEFVECLYNVDFDGAQKLRCEESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300

QY 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360

Db 301 CISINMLADKLNMTPEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQQVIEKTK 360

QY 361 SLSFRSOMLANNIEKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSOMLANNIEKLNQNSRSEAPNWTQDSGFY 396

RESULT 7

ABU05158

ID ABU05158 standard; protein; 396 AA.

XX AC ABU05158;

XX 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1824.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
PS Example 2; SEQ ID NO 1824; 134pp; English.
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDYKLYSDSDIPHALREKRTTVVAQLKQLQAEPIVKMFDPETTRMQSTRD 60
Db 1 MVDFAVDYKLYSDSDIPHALREKRTTVVAQLKQLQAEPIVKMFDPETTRMQSTRD 60
QY 61 GRMLFDYLDKRGFRQYEDLTLYRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
Db 61 GRMLFDYLDKRGFRQYEDLTLYRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
QY 121 WGLKASEILMNQWDAAMEDLTLKKTIDNNSSVSPLOSLOQTWLIHWSLFFVFNHPKGR 180
Db 121 WGLKASEILMNQWDAAMEDLTLKKTIDNNSSVSPLOSLOQTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQESYTYK 240
Db 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
QY 301 CISINMLADKLNMTPEERWTVNLIARNLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 360
Db 301 CISINMLADKLNMTPEERWTVNLIARNLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 360
QY 361 SLSPRSQMLANNIEKLNQSRSEAPNATQDSGFY 396
Db 361 SLSPRSQMLANNIEKLNQSRSEAPNATQDSGFY 396

RESULT 8
ADJ70258
ID ADJ70258 standard; protein; 396 AA.
XX
AC ADJ70258;
XX
DT 06-MAY-2004 (first entry)
DE Human heat mitochondrial protein as a therapeutic target SeqID2064.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;

KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003087768-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010870.
XX
PR 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
PA
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX WPI; 2003-845369/78.
XX
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
PS Claim 1; SEQ ID NO 2064; 180pp; English.
XX
CC This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nontropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
SQ Sequence 396 AA;
Query Match 100.0%; Score 2071; DB 7; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDYKLYSDSDIPHALREKRTTVVAQLKQLQAEPIVKMFDPETTRMQSTRD 60
Db 1 MVDFAVDYKLYSDSDIPHALREKRTTVVAQLKQLQAEPIVKMFDPETTRMQSTRD 60
QY 61 GRMLFDYLDKRGFRQYEDLTLYRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
Db 61 GRMLFDYLDKRGFRQYEDLTLYRYAKFYQECGNYSGAAEYLYFPRVLVPATDRNALSSL 120
QY 121 WGLKASEILMNQWDAAMEDLTLKKTIDNNSSVSPLOSLOQTWLIHWSLFFVFNHPKGR 180
Db 121 WGLKASEILMNQWDAAMEDLTLKKTIDNNSSVSPLOSLOQTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQESYTYK 240
Db 181 DNIIDFLYQOYLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQESYTYK 240
QY 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIFETFCRIHQ 300

```
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARLDAKIDSKLGHVVMGNNVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLRNARLDAKIDSKLGHVVMGNNVSPYQOVIETK 360
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 9
ADN30996
ID ADN30996 standard; protein; 396 AA.
XX
AC ADN30996;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human Int6 protein.
XX
KW Human; Int6; mammary epithelial cellular growth; cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN US6737251-B2.
XX
PD 18-MAY-2004.
XX
PF 14-MAY-2001; 2001US-00858152.
XX
PR 09-FEB-1995; 95US-00385998.
PR 09-FEB-1996; 96US-00875847.
PR 09-FEB-1996; 96WO-US001884.
PR 23-AUG-1999; 99US-00378842.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
DR WPI; 2004-387097/36.
XX
DR N-PSDB; ADN30995.
XX
PT Novel tumor Int6 recombinant protein that deregulates mammary epithelial
PT cellular growth, useful for treating cancer.
XX
PS Claim 1; SEQ ID NO 4; 44pp; English.
XX
CC The invention relates to the Int6 protein and the cDNA encoding it. The
CC Int6 protein deregulates mammary epithelial cellular growth. The cDNA and
CC protein are useful as vaccines for treating cancer. This sequence
CC represents the human Int6 protein of the invention.
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 8; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDYKKNLYSDDIPHALREKRTTVVAQLKQLQAEPIVKMFEDPETTRQMQSTRD 60
Db 1 MVDFAVDYKKNLYSDDIPHALREKRTTVVAQLKQLQAEPIVKMFEDPETTRQMQSTRD 60
QY 61 GRMLFDYLDKRGFRQEVLDTLRYAKFQYECGNYSGAAEYLFFRVLVPATDRNALSSL 120
Db 61 GRMLFDYLDKRGFRQEVLDTLRYAKFQYECGNYSGAAEYLFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAMEDLTRELKETIDNNSSVSPLOSQLOQTWLIHWSLFFVFNHPKGR 180
Db 121 WGLKASEILMQNWDAMEDLTRELKETIDNNSSVSPLOSQLOQTWLIHWSLFFVFNHPKGR 180
QY 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQOBSYTYK 240
Db 181 DNIIDFLYQPOLNAIQTMCPHILRYLTAVITNKDKVRKQVLDKLVKVIQOBSYTYK 240
```

```
QY 241 DPITEFVBCLYVNFDPDGAOKKLRECESVLVNDFFLVACLEDPIENARLFIETFCRIHQ 300
Db 241 DPITEFVBCLYVNFDPDGAOKKLRECESVLVNDFFLVACLEDPIENARLFIETFCRIHQ 300
QY 301 CISINMLADKLNMTPEEAERWIVNLRNARLDAKIDSKLGHVVMGNNVSPYQOVIETK 360
Db 301 CISINMLADKLNMTPEEAERWIVNLRNARLDAKIDSKLGHVVMGNNVSPYQOVIETK 360
QY 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 361 SLSFRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396

RESULT 10
ADN30994
ID ADN30994 standard; protein; 396 AA.
XX
AC ADN30994;
XX
DT 29-JUL-2004 (first entry)
XX
DE Murine Int6 protein.
XX
KW Mouse; Int6; mammary epithelial cellular growth; cancer; cytostatic.
XX
OS Mus musculus.
XX
PN US6737251-B2.
XX
PD 18-MAY-2004.
XX
PF 14-MAY-2001; 2001US-00858152.
XX
PR 09-FEB-1995; 95US-00385998.
PR 09-FEB-1996; 96US-00875847.
PR 09-FEB-1996; 96WO-US001884.
PR 23-AUG-1999; 99US-00378842.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Marchetti A, Buttitta F, Smith GH, Callahan R;
XX
DR WPI; 2004-387097/36.
XX
DR N-PSDB; ADN30993.
XX
PT Novel tumor Int6 recombinant protein that deregulates mammary epithelial
PT cellular growth, useful for treating cancer.
XX
PS Example 4; SEQ ID NO 2; 44pp; English.
XX
CC The invention relates to the Int6 protein and the cDNA encoding it. The
CC Int6 protein deregulates mammary epithelial cellular growth. The cDNA and
CC protein are useful as vaccines for treating cancer. This sequence
CC represents the murine Int6 protein of the invention.
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 2071; DB 8; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.2e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDYKKNLYSDDIPHALREKRTTVVAQLKQLQAEPIVKMFEDPETTRQMQSTRD 60
Db 1 MVDFAVDYKKNLYSDDIPHALREKRTTVVAQLKQLQAEPIVKMFEDPETTRQMQSTRD 60
QY 61 GRMLFDYLDKRGFRQEVLDTLRYAKFQYECGNYSGAAEYLFFRVLVPATDRNALSSL 120
Db 61 GRMLFDYLDKRGFRQEVLDTLRYAKFQYECGNYSGAAEYLFFRVLVPATDRNALSSL 120
QY 121 WGLKASEILMQNWDAMEDLTRELKETIDNNSSVSPLOSQLOQTWLIHWSLFFVFNHPKGR 180
Db 121 WGLKASEILMQNWDAMEDLTRELKETIDNNSSVSPLOSQLOQTWLIHWSLFFVFNHPKGR 180
```

QY	181	DNIIDLFLYQPOYLNAIQTMCPHILRYLT	TAVTINKDVRKRRQVLKDLVKVIOQESYTK	240
Db	181	DNIIDLFLYQPOYLNAIQTMCPHILRYLT	TAVTINKDVRKRRQVLKDLVKVIOQESYTK	240
QY	241	DPITFEVCELYYNFDFDGAQKLCESV	LVNDFLVACLEDFIENARLFIPTFCRIHQ	300
Db	241	DPITFEVCELYYNFDFDGAQKLCESV	LVNDFLVACLEDFIENARLFIPTFCRIHQ	300
QY	301	CISINNLADKLNMTPEEAERWIVNL	IRNARLDAKIDSKLGHVYVGNNAVSPYQQVIEKTK	360
Db	301	CISINNLADKLNMTPEEAERWIVNL	IRNARLDAKIDSKLGHVYVGNNAVSPYQQVIEKTK	360
QY	361	SLSFSRQMLAMNIEKKLNONSRSRAPN	WATQDSGFY	396
Db	361	SLSFSRQMLAMNIEKKLNONSRSRAPN	WATQDSGFY	396

RESULT 11	
ABB57306	
ID	ABB57306 standard; protein; 445 AA.
XX	
AC	ABB57306;
XX	
DT	07-MAR-2002 (first entry)
XX	
DE	Mouse ischaemic condition related protein sequence SEQ ID NO:856.
XX	
KW	Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW	vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX	
OS	Mus musculus.
XX	
PN	WO200188188-A2.
XX	
PD	22-NOV-2001.
XX	
PF	18-MAY-2001; 2001WO-JP004192.
XX	
PR	18-MAY-2000; 2000JP-00145977.
XX	
PA	(UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX	
PI	Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX	
DR	WPI; 2002-034733/04.
DR	N-PSDB; ABI99770.
XX	
PT	Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT	expression levels of particular genes defined in the specification or by
PT	determining the expression profile of a gene group comprising these
PT	genes.
XX	
PS	Claim 2; Page 2125-2127; 2690pp; English.

Best Local Similarity 100.0%; Pred. NO. 8.5e-193;	
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MVDPAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRMQSTRD 60
Db	50 MVDPAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAEETPIVKMFEDPETTRMQSTRD 109
Qy	61 GRMLFDYLDADKHGFRQBYDLTRYAKFYQECNGYSAAEYLYFFRVLPATDRLNALSSL 122
Db	110 GRMLFDYLDADKHGFRQBYDLTRYAKFYQECNGYSAAEYLYFFRVLPATDRLNALSSL 169
Qy	121 WGLKASSILMONWDAAEMEDTRLKETIDNNSSVSPLOQORTWLIHWSLVPFVFNHPKGR 180
Db	170 WGLKASSILMONWDAAEMEDTRLKETIDNNSSVSPLOQORTWLIHWSLVPFVFNHPKGR 229
Qy	181 DNIIDLFLYQPOLYNAIQTCPHILRLYTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 240
Db	230 DNIIDLFLYQPOLYNAIQTCPHILRLYTAVITNKDVRKRRQVLKDLVKVIOQESYTYK 289
Qy	241 DPITEFVECLYVNFDFDGAOKKLRECESVLVNDFPFLVACLEDFTENARLFIETFCRIHQ 300
Db	290 DPITEFVECLYVNFDFDGAOKKLRECESVLVNDFPFLVACLEDFTENARLFIETFCRIHQ 349
Qy	301 CISINMLADKLNMTPEEAERWIVNLIERNALDAKIDSKLGHVVMGNNAVSPYQVIEKTK 360
Db	350 CISINMLADKLNMTPEEAERWIVNLIERNALDAKIDSKLGHVVMGNNAVSPYQVIEKTK 409
Qy	361 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 396
Db	410 SLSFRSOMLANNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 12	
ABR319935	
ID	ABR319935 standard; protein; 445 AA.
XX	
AC	ABR319935;
XX	
DT	11-AUG-2003 (first entry)
XX	
DE	Human prostate selective polypeptide Pr327.
XX	
KW	Prostate; molecular marker; cancer; cytostatic; gene therapy; human.
XX	
OS	Homo sapiens.
XX	
PN	WO2003014298-A2.
XX	
PD	20-FEB-2003.
XX	
PF	02-AUG-2002; 2002WO-US024431.
XX	
PR	03-AUG-2001; 2001US-0309470P.
XX	
PR	30-OCT-2001; 2001US-0330747P.
XX	
PA	(ORIG-) ORIGENE TECHNOLOGIES INC.
XX	
PI	Sun Z, Li X, Jay G, Kovacs KF, Shu Y, Fan W;
XX	
XX	WPI; 2003-256562/25.
DR	N-PSDB; ACC47340.
XX	
PT	New polynucleotide, useful for preparing a composition for treating
PT	prostate disease, e.g., cancer.
XX	
PS	Claim 5; Page 150-152; 212pp; English.
XX	
CC	The invention relates to prostate selective polynucleotides and
CC	polypeptides. The polynucleotides are expressed in prostate and are
CC	useful as molecular markers, as drug targets, and for detecting,
CC	monitoring, preventing or treating diseases and conditions related to
CC	prostate, such as prostate cancers. The present sequence represents
CC	prostate specific polypeptide

```

XX      Sequence 445 AA;
SQ
Query Match      100.0%; Score 2071; DB 6; Length 445;
Best Local Similarity 100.0%; Pred. No. 8.5e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MYDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60
DB      50  MYDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 109

QY      61  GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSAAEYLYFFRVLPATDNRNALSSL 120
DB      110  GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSAAEYLYFFRVLPATDNRNALSSL 169

QY      121  WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 180
DB      170  WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 229

QY      181  DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDLVKVIQOESYTYK 240
DB      230  DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDLVKVIQOESYTYK 289

QY      241  DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB      290  DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 349

QY      301  CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSGLGHVVMGNNAVSPYQOVIETK 360
DB      350  CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSGLGHVVMGNNAVSPYQOVIETK 409

QY      361  SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
DB      410  SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 13
ABU05160
ID  ABU05160 standard; protein; 445 AA.
XX
AC  ABU05160;
XX
DT  29-JAN-2003 (first entry)
XX
DE  Human expressed protein tag (EPT) #1826.
XX
KW  Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW  protease; protease inhibitor; transporter; cytoskeletal protein;
KW  receptor; transcription factor; cancer; MHC;
KW  major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW  adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS  Homo sapiens.
XX
PN  WO200278524-A2.
XX
PD  10-OCT-2002.
XX
PF  28-MAR-2002; 2002WO-US009671.
XX
PR  28-MAR-2001; 2001US-0279495P.
PR  21-MAY-2001; 2001US-0292544P.
PR  08-AUG-2001; 2001US-0310801P.
PR  01-OCT-2001; 2001US-0326370P.
PR  04-DEC-2001; 2001US-0336780P.
PR  20-FEB-2002; 2002US-0358985P.
XX
PA  (ZYCO-) ZYCOS INC.
XX
PI  Chicx RM, Tomlinson AJ, Urban RG;
XX
DR  WPI; 2003-040607/03.
XX

```

```

PT      New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT      cytoskeletal proteins, receptors or transcription factors), useful for
PT      treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT      leukemia.
XX
XX      Example 2; SEQ ID NO 1826; 134pp; English.
XX
XX      The invention describes a purified polypeptide, which comprises a
XX      fragment of a kinase, phosphatase, protease, protease inhibitor,
XX      transporter, cytoskeletal protein, receptor or transcription factor. The
XX      polypeptide is useful as an immunogenic composition for eliciting in a
XX      mammal an immunogenic response directed against any of the purified
XX      polypeptide. The purified polypeptide, or the antibody that binds to this
XX      polypeptide, is useful for treating cancer. The polypeptide is also
XX      useful for identifying compounds that binds to a naturally processed
XX      class I or class II MHC-binding polypeptide. The polypeptides and
XX      polynucleotides are particularly useful for treating or preventing
XX      myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX      lymphoma or leukaemia. These are also useful for screening agents for
XX      treating the above mentioned diseases. This sequence represents an
XX      expressed protein tag (EPT) isolated from human tissue for translational
XX      profiling. Note: This sequence does not appear in the printed
XX      specification but was obtained in electronic format directly from WIPO at
XX      ftp.wipo.int/pub/published_pct_sequences
SQ      Sequence 445 AA;

Query Match      100.0%; Score 2071; DB 6; Length 445;
Best Local Similarity 100.0%; Pred. No. 8.5e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MYDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60
DB      50  MYDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 109

QY      61  GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSAAEYLYFFRVLPATDNRNALSSL 120
DB      110  GRMLFDYLADKHGFRQYELDTLYRYAKFOYECGNYSAAEYLYFFRVLPATDNRNALSSL 169

QY      121  WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 180
DB      170  WGLKASEILMQNWDAAEMDLTRLKETIDNNSVSSPLQSLQORTWLHWSLFFVFNHPKGR 229

QY      181  DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDLVKVIQOESYTYK 240
DB      230  DNIIDLFLYQOYLNAIQTMCPHILRYLTAVITNKDVRKRQVLDLVKVIQOESYTYK 289

QY      241  DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
DB      290  DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 349

QY      301  CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSGLGHVVMGNNAVSPYQOVIETK 360
DB      350  CISINMLADKLNMTPEAEERWIVNLIRNARLDKIDSGLGHVVMGNNAVSPYQOVIETK 409

QY      361  SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
DB      410  SLSPRSQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445

RESULT 14
ABU05163
ID  ABU05163 standard; protein; 445 AA.
XX
AC  ABU05163;
XX
DT  29-JAN-2003 (first entry)
XX
DE  Human expressed protein tag (EPT) #1829.
XX
KW  Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW  protease; protease inhibitor; transporter; cytoskeletal protein;
KW  receptor; transcription factor; cancer; MHC;

```

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX Homo sapiens.
 XX WO200278524-A2.
 XX PD 10-OCT-2002.
 XX PF 28-MAR-2002; 2002WO-US009671.
 XX PR 28-MAR-2001; 2001US-0279495P.
 XX PR 21-MAY-2001; 2001US-0292544P.
 XX PR 08-AUG-2001; 2001US-0310801P.
 XX PR 01-OCT-2001; 2001US-0326370P.
 XX PR 04-DEC-2001; 2001US-0336780P.
 XX PR 20-FEB-2002; 2002US-0358985P.
 XX PA (ZYCO-) ZYCOS INC.
 XX PI Chicx RM, Tomlinson AJ, Urban RG;
 XX WPI; 2003-040607/03.
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX Example 2; SEQ ID NO 1829; 134pp; English.
 CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 445 AA;
 Query Match 100.0%; Score 2071; DB 6; Length 445;
 Best Local Similarity 100.0%; Pred. No. 8.5e-193;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVDFAVDYKLYSDIPHALREKTTVVAQLKQLQAEPTVVKMFDPETTRQMSTRD 60
 DB 50 MVDFAVDYKLYSDIPHALREKTTVVAQLKQLQAEPTVVKMFDPETTRQMSTRD 109
 QY 61 GRMLFDYLDKRGFRQYLDLTLYRAKQFYECGNYSGAAEYLYFFRVLVDPATDRNALSSL 120
 DB 110 GRMLFDYLDKRGFRQYLDLTLYRAKQFYECGNYSGAAEYLYFFRVLVDPATDRNALSSL 169
 QY 121 WGLKASEILMQWDAAMEDLTFLKETIDNNSVSPLOSQORTWLHWSLFFVFNHPKGR 180
 DB 170 WGLKASEILMQWDAAMEDLTFLKETIDNNSVSPLOSQORTWLHWSLFFVFNHPKGR 229
 QY 181 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKDKRQVLDKLVKVIQQESYTYK 240
 DB 230 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKDKRQVLDKLVKVIQQESYTYK 289
 QY 241 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300
 DB 290 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 349
 QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARDAKIDSKLGHVVMGNNNAVSPYQVIEKTK 360
 DB 350 CISINMLADKLNMTPEEAERWIVNLIRNARDAKIDSKLGHVVMGNNNAVSPYQVIEKTK 409
 QY 361 SLSFRSQMLAWNIEKKLNQNSRSEAPNWTQDSGFY 396
 DB 410 SLSFRSQMLAWNIEKKLNQNSRSEAPNWTQDSGFY 445
 RESULT 15
 ABU05157
 ID ABU05157 standard; protein; 445 AA.
 XX AC ABU05157;
 XX DT 29-JAN-2003 (first entry)
 XX Human expressed protein tag (EPT) #1823.
 XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX Homo sapiens.
 XX WO200278524-A2.
 XX PD 10-OCT-2002.
 XX PF 28-MAR-2002; 2002WO-US009671.
 XX PR 28-MAR-2001; 2001US-0279495P.
 XX PR 21-MAY-2001; 2001US-0292544P.
 XX PR 08-AUG-2001; 2001US-0310801P.
 XX PR 01-OCT-2001; 2001US-0326370P.
 XX PR 04-DEC-2001; 2001US-0336780P.
 XX PR 20-FEB-2002; 2002US-0358985P.
 XX PA (ZYCO-) ZYCOS INC.
 XX PI Chicx RM, Tomlinson AJ, Urban RG;
 XX WPI; 2003-040607/03.
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX Example 2; SEQ ID NO 1829; 134pp; English.
 CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 445 AA;
 Query Match 100.0%; Score 2071; DB 6; Length 445;
 Best Local Similarity 100.0%; Pred. No. 8.5e-193;
 Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVDFAVDYKLYSDIPHALREKTTVVAQLKQLQAEPTVVKMFDPETTRQMSTRD 60
 DB 50 MVDFAVDYKLYSDIPHALREKTTVVAQLKQLQAEPTVVKMFDPETTRQMSTRD 109
 QY 61 GRMLFDYLDKRGFRQYLDLTLYRAKQFYECGNYSGAAEYLYFFRVLVDPATDRNALSSL 120
 DB 110 GRMLFDYLDKRGFRQYLDLTLYRAKQFYECGNYSGAAEYLYFFRVLVDPATDRNALSSL 169
 QY 121 WGLKASEILMQWDAAMEDLTFLKETIDNNSVSPLOSQORTWLHWSLFFVFNHPKGR 180
 DB 170 WGLKASEILMQWDAAMEDLTFLKETIDNNSVSPLOSQORTWLHWSLFFVFNHPKGR 229
 QY 181 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKDKRQVLDKLVKVIQQESYTYK 240
 DB 230 DNIIDFLYQPOLYNAIOTMCPHILRYLTAVITNKDKRQVLDKLVKVIQQESYTYK 289
 QY 241 DPITFEVCLYVNFDFDGAQKLRCESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300

```
SQ Sequence 445 AA;
Query Match      100.0%; Score 2071; DB 6; Length 445;
Best Local Similarity 100.0%; Pred. No. 8.5e-193;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 60
Db 50 MYDFAMDVYKNLYSDDIPHALREKRTTVVAQLKQLQAETEPVVKMFEDPETTRQMOSTRD 109
Qy 61 GRMLFDYLDKKGFRQYELDTLYRYAKFOYECGNYSAAEYLYFFRVLPATDRNALSSL 120
Db 110 GRMLFDYLDKKGFRQYELDTLYRYAKFOYECGNYSAAEYLYFFRVLPATDRNALSSL 169
Qy 121 WCKLASEIILMKNWDAAMEDLTRKETIDNNSVSSPLQSLQORTWLIHNSLFVFFNHPKGR 180
Db 170 WCKLASEIILMKNWDAAMEDLTRKETIDNNSVSSPLQSLQORTWLIHNSLFVFFNHPKGR 229
Qy 181 DNIIDLFLYQOVLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVIQOESYTYK 240
Db 230 DNIIDLFLYQOVLNAIQTMCPHILRYLTAVITNKDVRKRQVLKDLVKVIQOESYTYK 289
Qy 241 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 300
Db 290 DPITEFVECLYVNFDFDGAQKLRCESVLVNDFFLVACLEDFIENARLFIETFCRIHQ 349
Qy 301 CISINMLADKLNMTPEEAERWIWNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
Db 350 CISINMLADKLNMTPEEAERWIWNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409
Qy 361 SLSFRSOMLAMNIEKKLNONSSEAPFNWATQDSGFY 396
Db 410 SLSFRSOMLAMNIEKKLNONSSEAPFNWATQDSGFY 445
```

Search completed: March 8, 2005, 20:04:18
Job time : 167 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2005, 20:17:10 ; Search time 174 Seconds

(without alignments)
1165.422 Million cell updates/sec

Title: US-10-783-415-4

Perfect score: 396

Sequence: 1 MVDAMDYVKNLYSDIPHA.....LNQNSRSEAPNWTQDSGCFY 396

Scoring table: 40EIGGO

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_03:
1: uniprot_sprot;
2: uniprot_trembl;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	396	100.0	396	Q9BRV2	Q9brv2 homo sapien
2	396	100.0	445	1 IF36 HUMAN	P60228 homo sapien
3	396	100.0	445	1 IF36 MOUSE	P60229 mus musculus
4	396	100.0	445	2 Q641X8	Q641x8 rattus norv
5	373	94.2	422	2 Q9CT23	Q9ct23 mus musculus
6	295	74.5	445	2 Q8WVK4	Q8wvk4 homo sapien
7	295	74.5	445	2 Q6TAX5	Q6tax5 homo sapien
8	143	36.1	423	2 Q7ZXA5	Q7zxa5 xenopus lae
9	143	36.1	446	2 Q6P7L9	Q6p7l9 xenopus tro
10	105	26.5	119	2 Q8BNE6	Q8bne6 mus musculus
11	79	19.9	446	2 Q918W7	Q918w7 xenopus lae
12	73	18.4	73	2 Q6LCK2	Q6lck2 mus musculus
13	72	18.2	446	2 Q6DRI1	Q6dri1 brachydanio
14	72	18.2	446	2 Q6P7X8	Q6p7x8 brachydanio
15	72	18.2	448	2 Q6TH16	Q6th16 brachydanio
16	25	6.3	462	2 Q7QIL8	Q7qil8 anopheles g
17	15	3.8	418	2 Q9M2L8	Q9m2l8 arabidopsis
18	15	3.8	441	2 Q9C5Z3	Q9c5z3 arabidopsis
19	15	3.8	441	2 Q9M4T7	Q9m4t7 arabidopsis
20	13	3.3	372	2 Q8MR88	Q8mr88 drosophila
21	13	3.3	435	1 IF36 DROME	Q77410 drosophila
22	10	2.5	397	2 Q69W37	Q69w37 oryza sativ
23	10	2.5	415	2 Q69W36	Q69w36 oryza sativ
24	10	2.5	439	2 Q8GV10	Q8gv10 oryza sativ
25	9	2.3	421	2 Q8CQ02	Q8cq02 yarrowia li
26	9	2.3	517	2 Q81315	Q81315 plasmodium
27	9	2.3	527	2 Q7RQ22	Q7rq22 plasmodium
28	8	2.0	136	2 Q7LZ00	Q7lz00 ovine lenti
29	8	2.0	141	2 Q841L1	Q841l1 streptomyce
30	8	2.0	175	1 SSRP PROMP	Q7uzp2 prochloroco
31	8	2.0	202	2 Q6DFE2	Q6dfe2 xenopus lae

32 8 2.0 215 2 Q6D0T5 Q6d0t5 erwinia car
33 8 2.0 299 2 Q6FB96 Q6fb96 acinetobact
34 8 2.0 335 1 COBT_TRETH Q781c7 thermus the
35 8 2.0 336 2 Q746P7 Q746p7 thermus the
36 8 2.0 368 2 Q6Y686 Q6y686 mus musculus
37 8 2.0 393 2 Q9PLP1 Q9plp1 chlamydia m
38 8 2.0 395 2 Q6Y687 Q6y687 homo sapien
39 8 2.0 425 2 Q6D0T5 Q6d0t5 erwinia car
40 8 2.0 425 2 Q6D2J4 Q6d2j4 erwinia car
41 8 2.0 432 1 IF36 CAEEL Q61820 caenorhabd
42 8 2.0 432 2 Q9Q8W7 Q9q8w7 rabbit fibr
43 8 2.0 436 2 Q6Z475 Q6z475 oryza sativ
44 8 2.0 436 2 Q9LNL2 Q9lnl2 arabidopsis
45 8 2.0 493 1 AMPA_CLOPE Q8xhi3 clostridium

ALIGNMENTS

RESULT 1

Q9BRV2 PRELIMINARY; PRT; 396 AA.
AC Q9BRV2;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE EIF356 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RT Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005944; AAH05944.1; -
DR InterPro; IPR001717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 396 AA; 46445 MW; 09165D0D612B1C3B CRC64;

Query Match 100.0%; Score 396; DB 2; Length 396;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDAMDYVKNLYSDIPHA...LNQNSRSEAPNWTQDSGCFY 396

|||||

Db 1 MVDPMADYKXLYSDDIPHALREKRTTVAOLKQQAETEPVVKQFDEPTTROMQSTRD 60
Qy 61 GRMLFDYLADKHGFRQEQYDLTYAKFOYECGNSGAAEYLYFRVLVLPATDNLSSSL 120
Db 61 GRMLFDYLADKHGFRQEQYDLTYAKFOYECGNSGAAEYLYFRVLVLPATDNLSSSL 120
Qy 121 WGLKASEILMQLMWDAAEDLRLKLTETDNNSSVSPLOSLQORTWLHWSLFPVFNHPKGR 180
Db 121 WGLKASEILMQLMWDAAEDLRLKLTETDNNSSVSPLOSLQORTWLHWSLFPVFNHPKGR 180
Qy 181 DNIIDLFLYQOYLNAITQCMTPCHILRLTYTAVITNKDVRKROVLKDLVKVIOQESYTK 240
Db 181 DNIIDLFLYQOYLNAITQCMTPCHILRLTYTAVITNKDVRKROVLKDLVKVIOQESYTK 240
Qy 241 DPITEFVECLVNFDFDGAQKLRCECVLNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Db 241 DPITEFVECLVNFDFDGAQKLRCECVLNDFFLVACLEDFTENARLFIFETFCRIHQ 300
Qy 301 CINSIMLADKLMTPEEAERIVNLIRNARDKIDSKLGHVVMGNNAVSPYQVIEKTK 360
Db 301 CINSIMLADKLMTPEEAERIVNLIRNARDKIDSKLGHVVMGNNAVSPYQVIEKTK 360
Qy 361 SLSFRSOMLAWNIEKLNQNSRSEAPNATQDSGFY 396
Db 361 SLSFRSOMLAWNIEKLNQNSRSEAPNATQDSGFY 396

RESULT 2

IF36_HUMAN
ID IF36_HUMAN STANDARD; PRT; 445 AA.
AC P60228; O43902; Q64058; Q64059; Q64252;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Eukaryotic translation initiation factor 3 subunit 6 (eIF-3 p48)
GN (eIF3e) (Viral integration site protein INT-6 homolog).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 280-289 AND 427-436.
RC TISSUE=Liver;
RX MEDLINE=97442403; PubMed=9295280; DOI=10.1074/jbc.272.38.23477;
RA Asano K., Merrick W.C., Hershey J.W.B.;
RT "The translation initiation factor eIF3-p48 subunit is encoded by int-6, a site of frequent integration by the mouse mammary tumor virus genome.";
RT J. Biol. Chem. 272:23477-23480 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=98066777; PubMed=9403073; DOI=10.1006/geno.1997.4996;
RA Miyazaki S., Imatani A., Ballard L., Marchetti A., Buttitta F.,
RA Albertsen H., Nevanlinna H.A., Gallahan D., Callahan R.;
RT "The chromosome location of the human homolog of the mouse mammary tumor-associated gene INT6 and its status in human breast carcinomas.";
RT Genomics 46:155-158 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96337990; PubMed=8688078;
RA Desbois C., Rousset R., Bantignies F., Jalinot P.;
RT "Exclusion of int-6 from PML nuclear bodies by binding to the HTLV-I Tax oncoprotein.";
RT Science 273:951-953 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=12386384;
RA Neuvirt C., Jin D.-Y., Semmes O.J., Diella F., Callahan R.,
RA Jeang K.-T.;
RT "Divergent subcellular locations of HTLV-I Tax and Int-6: a contrast

between in vitro protein-protein binding and intracellular protein colocalization.";
J. Biomed. Sci. 4:229-234 (1997).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow, Brain, and Muscle;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausberg R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [6]
RP INTERACTION WITH TRIM27.
RX PubMed=10504338;
RA Morris-Desbois C., Bochar V., Reynaud C., Jalinot P.;
RT "Interaction between the Ret finger protein and the int-6 gene product and co-localisation into nuclear bodies.";
RL J. Cell Sci. 112:3331-3342 (1999).
CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of methionyl-tRNAi and mRNA.
CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits. Interacts with TRIM27.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- SIMILARITY: Belongs to the eIF3S6 family.
CC -!- SIMILARITY: Contains 1 PCI domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U54562; AAC51760.1; -
CC EMBL; U94174; AAC51917.1; -
CC EMBL; U94162; AAC51917.1; JOINED.
CC EMBL; U94163; AAC51917.1; JOINED.
CC EMBL; U94164; AAC51917.1; JOINED.
CC EMBL; U94165; AAC51917.1; JOINED.
CC EMBL; U94166; AAC51917.1; JOINED.
CC EMBL; U94167; AAC51917.1; JOINED.
CC EMBL; U94168; AAC51917.1; JOINED.
CC EMBL; U94169; AAC51917.1; JOINED.
CC EMBL; U94170; AAC51917.1; JOINED.
CC EMBL; U94171; AAC51917.1; JOINED.
CC EMBL; U94172; AAC51917.1; JOINED.
CC EMBL; U94173; AAC51917.1; JOINED.
CC EMBL; U94175; AAC51919.1; -
CC EMBL; U62962; AAB58251.1; -
CC EMBL; U85947; AAB88873.1; -
CC EMBL; BC000734; AAH00734.1; -
CC EMBL; BC008419; AAH08419.1; -
CC EMBL; BC016706; AAH16706.1; -
CC EMBL; BC021679; AAH21679.1; -
CC IntAct; P60228; -

Genew; HGNC:3277; EIF3S6.
H-InvDB; HIX0007722; -.
DR Reactome; P60228; -.
DR MIM; 602210; -.
DR InterPro; IPR000717; PCI.
KW Direct protein sequencing; Initiation factor; Protein biosynthesis.
SQ SEQUENCE 445 AA; 52220 MW; A5368651DDDDDDOC CRC64;
Query Match 100.0%; Score 396; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVDFAVDVYKNLYSDDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPTTRQMOSTRD 60
Db 50 MVDFAVDVYKNLYSDDDIPHALREKRTTVAQLKQLQAEETPIVKMFEDPTTRQMOSTRD 109
Qy 61 GRMLFDYLDKHGFRQBYLDTLYRYAKFOYECGNYSGAAYLYFFRVLVPATRNALSSL 120
Db 110 GRMLFDYLDKHGFRQBYLDTLYRYAKFOYECGNYSGAAYLYFFRVLVPATRNALSSL 169
Qy 121 WGLASBILMNQNDAAEMEDTLKKEITIDNNSSVPSLQSLQORTWLHWSLFFVFNHPKGR 180
Db 170 WGLASBILMNQNDAAEMEDTLKKEITIDNNSSVPSLQSLQORTWLHWSLFFVFNHPKGR 229
Qy 181 DNIIDLFLYQPYLNAIQTCWPHILRYLTAVITNKDVRKRRQVLDLVKVIQOESYTYK 240
Db 230 DNIIDLFLYQPYLNAIQTCWPHILRYLTAVITNKDVRKRRQVLDLVKVIQOESYTYK 289
Qy 241 DPITEFVECLVNVDFDGAOKKLCESVNVDFLVLACLEDFTIENARLFIETFCRIHQ 300
Db 290 DPITEFVECLVNVDFDGAOKKLCESVNVDFLVLACLEDFTIENARLFIETFCRIHQ 349
Qy 301 CISINMLADKLNMTPEAEERWVNLINRNLDAKIDSKLGHVVMGNNAVSPYQOVIETKT 360
Db 350 CISINMLADKLNMTPEAEERWVNLINRNLDAKIDSKLGHVVMGNNAVSPYQOVIETKT 409
Qy 361 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 396
Db 410 SLFSRQMLAMNIEKKLNQNSRSEAPNWTQDSGFY 445
RESULT 3
ID IF36 MOUSE STANDARD; PRT; 445 AA.
AC P60229; Q43902; Q64058; Q64252;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Eukaryotic translation initiation factor 3 subunit 6 (EIF3 p48)
DE (EIF3) (Mammary tumor-associated protein INT-6) (Viral integration site protein INT-6) (MWTV integration site 6).
GN Names=EIF3e6; Synonyms=Int6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95156630; PubMed=7853537;
RA Marchetti A.; Buttitta F.; Miyazaki S.; Gallahan D.; Smith G.H.; Callahan R.;
RT "Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia.";
RN J. Virol. 69:1932-1938 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97405883; PubMed=9260927;
RA Diella F.; Levi G.; Callahan R.;
RT "Characterization of the INT6 mammary tumor gene product.";
RN DNA Cell Biol. 16:839-847 (1997).
RN [3]
RP REVISIONS TO N-TERMINUS.
RA Callahan R.;

```

Qy 121 WGLASEILMNQWDAAMEDLTRLKETIDNNSSVSPLOSLQOQRTWLIHWSLFFVFNHPKGR 180
Db 170 WGLASEILMNQWDAAMEDLTRLKETIDNNSSVSPLOSLQOQRTWLIHWSLFFVFNHPKGR 229
Qy 181 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240
Db 230 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 289
Qy 241 DPITEFVECLVNFDFDGAQKKLRECESVLVNDFLVLACLEDFFENARLFIPTFCRIHQ 300
Db 290 DPITEFVECLVNFDFDGAQKKLRECESVLVNDFLVLACLEDFFENARLFIPTFCRIHQ 349
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 360
Db 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 409
Qy 361 SLSPRSQMLANNIEKKLNQNSRSEAPNWATQDSGFY 396
Db 410 SLSPRSQMLANNIEKKLNQNSRSEAPNWATQDSGFY 445

RESULT 4
Q641X8 PRELIMINARY; PRT; 445 AA.
AC Q641X8;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein. (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082087; AA082087.1; -.
KW Hypothetical protein.
SQ SEQUENCE 445 AA; 52220 MW; AS368651DD0DD0C CRC64;
Query Match 100.0%; Score 396; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVDFAVDYKYNLYSDSDIPHALREKRTTVAQKQLQAEITEPIVKMFEDPETTRQMQSTRD 60
Db 50 MVDFAVDYKYNLYSDSDIPHALREKRTTVAQKQLQAEITEPIVKMFEDPETTRQMQSTRD 109

```

```

Qy 61 GRMLFDYLADKHGFRQEYLDLTRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 120
Db 110 GRMLFDYLADKHGFRQEYLDLTRYAKFOYECGNYSGAAEYLYFFRVLPATDRNALSSL 169
Qy 121 WGLASEILMNQWDAAMEDLTRLKETIDNNSSVSPLOSLQOQRTWLIHWSLFFVFNHPKGR 180
Db 170 WGLASEILMNQWDAAMEDLTRLKETIDNNSSVSPLOSLQOQRTWLIHWSLFFVFNHPKGR 229
Qy 181 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 240
Db 230 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDVRKROVLKDLVKVIOQESYTYK 289
Qy 241 DPITEFVECLVNFDFDGAQKKLRECESVLVNDFLVLACLEDFFENARLFIPTFCRIHQ 300
Db 290 DPITEFVECLVNFDFDGAQKKLRECESVLVNDFLVLACLEDFFENARLFIPTFCRIHQ 349
Qy 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 360
Db 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQVIEKTK 409
Qy 361 SLSPRSQMLANNIEKKLNQNSRSEAPNWATQDSGFY 396
Db 410 SLSPRSQMLANNIEKKLNQNSRSEAPNWATQDSGFY 445

RESULT 5
Q9CT23 PRELIMINARY; PRT; 422 AA.
AC Q9CT23;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2610017H09 product:mammary tumor integration
DE site 6, full insert sequence. (Fragment).
GN Name=Bif386;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=93279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RN [5]

```

SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (KISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Hanagaki T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoaka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK011436; BAB27621.1; -;
DR MGD; MGI:99257; Eif3a6.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00089; PINT; 1.
FT NON TER 422 422
SQ SEQUENCE 422 AA; 49568 MW; A05157DBCC6D375B CRC64;

Query Match 94.2%; Score 373; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVDFAVDYKNLYSDIPHALREKRTTVAQLKQLQAEPIVKMFDPETTRQMOSTRD 60
DB 50 MVDFAVDYKNLYSDIPHALREKRTTVAQLKQLQAEPIVKMFDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKXGFRQYEDLTLYRYAKFOYECGNTSGAAEYLYFFRVLVPAIDRNALSSL 120
DB 110 GRMLFDYLDKXGFRQYEDLTLYRYAKFOYECGNTSGAAEYLYFFRVLVPAIDRNALSSL 169
QY 121 WGLKASEILMQNDAAEDLTFLKKTIDNNSSVSSPLQSLQRTWLHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMQNDAAEDLTFLKKTIDNNSSVSSPLQSLQRTWLHWSLFFVFNHPKGR 229
QY 181 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 240
DB 230 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 289
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 349
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLFSRQMLAMNI 373
DB 410 SLFSRQMLAMNI 422

RESULT 6
Q8WVK4 PRELIMINARY; PRT; 445 AA.
ID Q8WVK4

Q8WVK4;
AC 01-WAR-2002 (TREMELrel. 20, Created)
DT 01-WAR-2002 (TREMELrel. 20, Last sequence update)
DE 01-WAR-2004 (TREMELrel. 26, Last annotation update)
DE Mammary tumor integration site 6 (Oncogene homolog).
GN Name=EIP356;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017887; AAH17887.1; -;
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00089; PINT; 1.
SQ SEQUENCE 445 AA; 52248 MW; AFOAD651DAAA8D0C CRC64;

Query Match 74.5%; Score 295; DB 2; Length 445;
Best Local Similarity 99.7%; Pred. No. 1.4e-299;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MVDFAVDYKNLYSDIPHALREKRTTVAQLKQLQAEPIVKMFDPETTRQMOSTRD 60
DB 50 MVDFAVDYKNLYSDIPHALREKRTTVAQLKQLQAEPIVKMFDPETTRQMOSTRD 109
QY 61 GRMLFDYLDKXGFRQYEDLTLYRYAKFOYECGNTSGAAEYLYFFRVLVPAIDRNALSSL 120
DB 110 GRMLFDYLDKXGFRQYEDLTLYRYAKFOYECGNTSGAAEYLYFFRVLVPAIDRNALSSL 169
QY 121 WGLKASEILMQNDAAEDLTFLKKTIDNNSSVSSPLQSLQRTWLHWSLFFVFNHPKGR 180
DB 170 WGLKASEILMQNDAAEDLTFLKKTIDNNSSVSSPLQSLQRTWLHWSLFFVFNHPKGR 229
QY 181 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 240
DB 230 DNIIDFLYQOYLNAIOTMCPHILRYLTAVITNKDKVRKROVLKDLVKVIOQESYTYK 289
QY 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 300
DB 290 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIPTFCRIHQ 349
QY 301 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 360
DB 350 CISINMLADKLNMTPEEAERWIVNLIRNARLDAKIDSKLGHVVMGNNAVSPYQOVIETK 409
QY 361 SLFSRQMLAMNI 422

```
|||||
410 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 445

RESULT 7
Q61AX5
ID Q61AX5 PRELIMINARY; PRT; 445 AA.
AC Q61AX5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE EIF36 protein.
GN Name=EIF36;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OK NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR457029; CAG33310.1; -.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 445 AA; 52205 MW; C7C0PF755BD9A8D4 CRC64;

Query Match 74.5%; Score 295; DB 2; Length 445;
Best Local Similarity 99.7%; Pred. No. 1.4e-299;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MVDFAVDYKKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFDPETTRQMQSTRD 60
Db 50 MVDFAVDYKKNLYSDDIPHALREKRTTVVAQLKQLAETEPVVKMFDPETTRQMQSTRD 109
Qy 61 GRMLFDYLADKHGRFQEVLDLTLYAKFQYECGNYSGAAEYLYFRVLVPAVDNALSSL 120
Db 110 GRMLFDYLADKHGRFQEVLDLTLYAKFQYECGNYSGAAEYLYFRVLVPAVDNALSSL 169
Qy 121 WGLKASEILMNQWDAAMEDLRLKETIDNNSSVSPQSLQLOQTWLIHWSLVFFNHPKGR 180
Db 170 WGLKASEILMNQWDAAMEDLRLKETIDNNSSVSPQSLQLOQTWLIHWSLVFFNHPKGR 229
Qy 181 DNIIDLFLYQOYLNAIOTMCPHILRYLTAVITNKDVKRQVLKLVKVIQESYTYK 240
Db 230 DNIIDLFLYQOYLNAIOTMCPHILRYLTAVITNKDVKRQVLKLVKVIQESYTYK 289
Qy 241 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 300
Db 290 DPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLEDFIENARLFIFETFCRIHQ 349
Qy 301 CISINMLADKLNMTPEEAERIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQVIEKTK 360
Db 350 CISINMLADKLNMTPEEAERIVNLIRNARLDKIDSKLGHVVMGNNAVSPYQVIEKTK 409
Qy 361 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 396
Db 410 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 445

RESULT 8
Q7ZXAS
ID Q7ZXAS PRELIMINARY; PRT; 423 AA.
AC Q7ZXAS;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LOC398503 protein (fragment).
GN Name=LOC398503;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

|||||
410 SLSPRSQMLAMNIEKLNQNSRSEAPNWTQDSGFY 445
```

```
Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045079; AAH45079.1; -.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
FT NON TER
SQ SEQUENCE 423 AA; 49523 MW; F691314F43EF400B CRC64;

Query Match 36.1%; Score 143; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 162 RTWLIHWSLVFFNHPKGRDNIIDLFLYQOYLNAIOTMCPHILRYLTAVITNKDVKR 221
Db 189 RTWLIHWSLVFFNHPKGRDNIIDLFLYQOYLNAIOTMCPHILRYLTAVITNKDVKR 248
Qy 222 RQVLKDLVKVLIQESYTYKDPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLE 281
Db 249 RQVLKDLVKVLIQESYTYKDPITEFVECLYVNFDFDGAQKLRCEESVLVNDFLVACLE 308
Qy 282 DFENARLFIFETFCRIHQCSI 304
Db 309 DFENARLFIFETFCRIHQCSI 331

RESULT 9
Q6P7L9
ID Q6P7L9 PRELIMINARY; PRT; 446 AA.
AC Q6P7L9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 3, subunit 6..
```


GN Name=eif3-p48-prov;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Fipodea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8364;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=2238957; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altchul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBJ databases.
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC061611; AAH61611.1;
 DR GO; GO:0003743; P:translational initiation factor activity; IEA.
 DR InterPro; IPR000717; PCI.
 DR Pfam; PF01399; PCI; 1.
 DR SMART; SM00088; PINT; 1.
 KW Initiation factor.
 SQ SEQUENCE 446 AA; 52296 MW; EB12CE70888DA8B1 CRC64;
 Query Match 36.1%; Score 143; DB 2; Length 446;
 Best Local Similarity 100.0%; Pred. No. 2e-140;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 162 RTWLTHWSLFVFNHPKGRDNIIDLFLYQPVYLNAIQTMCPHLRYLTAVITNKDVRK 221
 DB 212 RTWLTHWSLFVFNHPKGRDNIIDLFLYQPVYLNAIQTMCPHLRYLTAVITNKDVRK 271
 QY 222 RQVLKDLVKVIOQESYTYKDPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLE 281
 DB 272 RQVLKDLVKVIOQESYTYKDPITEFVECLYVNFDFGAKKLRECESVLVNDFFLVACLE 331
 QY 282 DFENARLFIFFTFCRIHQICSI 304
 DB 332 DFENARLFIFFTFCRIHQICSI 354
 RESULT 10
 QBNE6
 ID QBNE6 PRELIMINARY; PRT; 119 AA.
 AC QBNE6;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
 DE enriched library, clone:D130036L03 product:mammary tumor integration
 DE site 6, full insert sequence.
 GN Name=Eif3e6;
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
 RA The FANTOM Consortium;
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN (5)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
 RX MEDLINE=20530913; PubMed=11078661; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Alizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN (6)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
 RA Adachi J., Alizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu M., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK083874; BAC39045.1;
 DR MGI; MGI:99257; Bif3e6.
 DR InterPro; IPR010935; SMC hinge.
 SQ SEQUENCE 119 AA; 14113 MW; ACC86F6ABBE052E6 CRC64;
 Query Match 26.5%; Score 105; DB 2; Length 119;
 Best Local Similarity 100.0%; Pred. No. 3.5e-101;
 Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 130 MQNDAAWEDLTRKETIDNNSSVSLQSLQRTWLHWSLFVFNHPKGRDNIIDLFLY 189

```
Db 1 MONWDAAMEDLTRLKETIDNNSSVSPLOSQLOQTWLIHWSLFPVFNHFKGRDNIIDFLY 60
Qy 190 QPYLNALQTCPHILRLYLTAVITNKDVRKRQVLKDLVKVIOQ 234
Db 61 QPYLNALQTCPHILRLYLTAVITNKDVRKRQVLKDLVKVIOQ 105

RESULT 11
Q918W7 PRELIMINARY; PRT; 446 AA.
AC Q918W7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Int-6 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Morris-Desbois C., Bochar V., Reynaud C., Jalinot P.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162775; AAF80474.1; -.
DR InterPro; IPR000717; PCI_hinge.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 446 AA; 52240 MW; 21C5605DA9CF88BF CRC64;

Query Match 19.9%; Score 79; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 2e-73;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 181 DNIIDFLYQPYLNALQTCPHILRLYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 240
Db 231 DNIIDFLYQPYLNALQTCPHILRLYLTAVITNKDVRKRQVLKDLVKVIOQESYTYK 290

Qy 241 DPITEFVECLYVNFDFGA 259
Db 291 DPITEFVECLYVNFDFGA 309

RESULT 12
Q6LCK2 PRELIMINARY; PRT; 73 AA.
AC Q6LCK2;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Int-6 protein (Fragment).
GN Name=Int-6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97442403; PubMed=9295280; DOI=10.1074/jbc.272.38.23477;
RT Asano K., Merrick W.C., Hershey J.W.;
RT "The translation initiation factor eIF3-p48 subunit is encoded by int-6, a site of frequent integration by the mouse mammary tumor virus genome.";
RT J. Biol. Chem. 272:23477-23480 (1997).
RL EMBL; U54563; AAC53346.1; -.
FT NON TER 73
SQ SEQUENCE 73 AA; 8565 MW; 0904AD1E79842124 CRC64;

Query Match 18.4%; Score 73; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 7.2e-68;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 FEDPETTRQMOSTRDGRMLFDYLDKKGFRQYLDTRYAKFYECGNGSAAEYLYFF 105
Db 96 FEDPETTRQMOSTRDGRMLFDYLDKKGFRQYLDTRYAKFYECGNGSAAEYLYFF 155

Qy 106 RVLVPATDRNAL 117
Db 156 RVLVPATDRNAL 167

RESULT 14
Q6P7X8 PRELIMINARY; PRT; 446 AA.
AC Q6P7X8;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 3, subunit 6.
GN ORFNames=zgc:63821;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
```


RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=whole body;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061454; AAH61454.1; -
DR ZFIN; ZDB-GENE-030131-3827; zgc:63821.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
KW Initiation factor.
SQ SEQUENCE 446 AA; 52389 MW; 32F9D000E69F4F87 CRC64;

Query Match 18.2%; Score 72; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 4.3e-66;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 FEDPTTQMOSTRGMFLFDYLDKKGPRQRYLDTLYRYAKFOYECGNYSGAAEYLYFF 105
DB 96 FEDPTTQMOSTRGMFLFDYLDKKGPRQRYLDTLYRYAKFOYECGNYSGAAEYLYFF 155

QY 106 RVLVPATDRNAL 117
DB 156 RVLVPATDRNAL 167

RESULT 15
Q6TH16
ID Q6TH16 PRELIMINARY; PRT; 448 AA.
AC Q6TH16;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 3, subunit 6 48kDa.
GN Name=EIP356; ORFNames=zgc:63821;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney marrow;
RA Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,
RA Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zou L.I.,
RA Kanki J.P., Look A.T., Chen Z.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY398340; AAQ97773.1; -
DR ZFIN; ZDB-GENE-030131-3827; zgc:63821.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR010935; SMC_hinge.
DR Pfam; PF01399; PCI; 1.

DR SMART; SM00088; PINT; 1.
KW Initiation factor.
SQ SEQUENCE 448 AA; 52672 MW; 190C9FD4CEB77803 CRC64;

Query Match 18.2%; Score 72; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 4.3e-66;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 LFLYQPYLNAIQTMCPHILRYLTAVITNKQVRKQVLDKLVKVIQOESYTYKDPITE 245
DB 238 LFLYQPYLNAIQTMCPHILRYLTAVITNKQVRKQVLDKLVKVIQOESYTYKDPITE 297

QY 246 FVECLYVNFDFD 257
DB 298 FVECLYVNFDFD 309

Search completed: March 8, 2005, 20:26:35
Job time : 177 secs

THIS PAGE BLANK (USPTO)